Development of tools for automated collection, integration and analysis of genetic data in ALS

Abel, Olubunmi

Awarding institution:
King's College London

DEVELOPMENT OF TOOLS FOR AUTOMATED COLLECTION, INTEGRATION AND ANALYSIS OF GENETIC DATA IN ALS

Olubunmi Kemi Deborah ABEL
PhD

Department of Clinical Neuroscience

Institute of Psychiatry

King's College London

Thesis submitted for the degree of Doctor of Philosophy

March 2014

Abstract

Amyotrophic Lateral Sclerosis (ALS), also known as Lou Gehrig's disease, typically leads to death within 3-5 years of symptom onset. Understanding what causes ALS has been a challenge, but more research in this area, enhanced by advanced technology like high-throughput next generation sequencing, is paving the way for better information and direction. The volume of data generated by genetics researchers has dramatically increased, largely because of increased opportunities for collaboration. ALSoD, a widely used online genetics database for collating, analysing and integrating ALS data, has been updated with analytics tools and is able to portray the data graphically to users. Mutations and other gene variants have been mapped to genomic coordinates, and the inclusion of dbSNP ids has been implemented to facilitate the integration of data from numerous public sources. To increase the usability and functionality of ALSoD, population frequency of each variant found in the 1000 Genome Project and Exome Variation Server (EVS) databases is displayed. To contribute to a better understanding of the pathogenesis of ALS, links to information on animal models are also available. Furthermore, ALSoD can now be viewed on mobile devices and for Android platforms a mobile app is also available.

Table Of Contents

Abstract 2
Table Of Contents 3
Table of Figures 10
Table of Tables 16
Table of Equations 16
Document Information 16
Acknowledgements 17
Chapter 1 INTRODUCTION AND BACKGROUND 19
1.1 Health is wealth 19
1.2 Diseases and Disorders 19
1.3 Motor Neuron Disease 19
1.4 Amyotrophic Lateral Sclerosis 20
1.5 Clinical representation of ALS 20
1.6 Diagnosis 21
1.7 Prevention and Cure 21
1.8 Clinical Trials 22
1.9 Palliative care / Equipment 22
1.10 Epistasis 22
1.11 Epigenetics 23
1.12 Epidemiology 23
1.13 Neuroimaging 23
1.14 In Vivo and in Vitro Models 23
1.15 Genetics of Amyotrophic Lateral Sclerosis 24
1.16 Environmental causes 25
1.17 How credible are these reported genes? 25
1.18 Origin of ALS-linked loci 26
1.19 Databases 52
1.19.1 Data 52
1.19.2 Evolution of Database 53

1.19.3 Relational model: review 53
1.19.4 Database Management Systems 53
1.19.5 Examples of Databases 53
1.20 Genetics database 53
1.20.1 History of Genetics Databases 53
1.20.2 Non-locus specific mutation databases 54
1.20.3 ALS-specific mutation databases 58
1.21 Collection 59
1.21.1 Data mining, Text mining and Web mining 59
1.21.2 External Databases 59
1.21.3 The need for a Single Resource 60
1.22 Integration 60
1.22.1 What is needed in ALS 60
1.22.2 Curators/Contributors 61
1.22.3 Google API 61
1.22.4 Animal Models 61
1.23 Analysis 61
1.23.1 Genome Wide Association Study (GWAS) 62
1.23.2 On-the-fly Meta Analysis 62
1.23.3 Predicting gene interactions 62
1.24 Smartphone Development 63
1.24.1 Mobile Websites 64
1.24.2 Apps 64
1.24.3 Advantages of a mobile website over an app 66
1.24.4 Advantages of an app over a mobile website 67
1.25 The Project 68
1.25.1 Problem Statement 68
1.25.2 Research Questions 68
1.25.3 Overall aim 69
1.25.4 Objectives 69
Chapter 2 Publication 1 - ALSoD: A User-Friendly Online Bioinformatics Tool for Amyotrophic Lateral Sclerosis Genetics [521] 70

Chapter 3 Publication 2 - Keeping up with genetic discoveries in amyotrophic lateral sclerosis: The ALSoD and ALSGene databases [486] 78
Chapter 4 Publication 3 - Credibility Analysis of Putative Disease-Causing Genes Using Bioinformatics[522] 91
Chapter 5 Publication 4 - Development of a Smartphone App for a Genetics Website: The Amyotrophic Lateral Sclerosis Online Genetics Data [523] 98
Chapter 6 METHODS 109
6.1 Research questions with an overview of proposed methods 109
6.2 Keeping up with genetic discoveries in amyotrophic lateral sclerosis: 110
6.2.1 Data Submission 110
6.2.2 Submit mutation data 111
6.2.3 Submit new gene 118
6.2.4 Submit patient data 120
6.2.5 Submit replicated mutation data 126
6.2.6 Submission process into ALSoD 127
6.2.7 Search terms on search engines 129
6.2.8 Contributors and Collaborators 130
6.2.9 Unique Identifiers 130
6.2.10 Graphical display of phenotypic data 130
6.2.11 Mapping variants on Google Earth 131
6.2.12 Meta-analysis and on-the-fly analysis of association data 131
6.2.13 Instruction Manual on updating database 131
6.2.14 Updating Database 131
6.3 ALSoD: A user-friendly online bioinformatics tool for amyotrophic lateral sclerosis Genetics 135
6.3.1 Funding and Sponsorship 135
6.3.2 Programming 140
6.3.3 Web Design 143
6.3.3.2 MasterPage 150
6.3.3.3 Homepage 160
6.3.4 On-the-fly Meta-Analysis of Genome Wide Association Studies in ALS 173
6.3.5 Collaborations 187
6.3.6 Embedded bioinformatics tools 187

6.3.7 Imported bioinformatics data 188
6.3.8 Population Frequency from 1000genome and EVS 191
6.3.9 Animal Models 194
6.3.10 Integrated Bioinformatics links 197
6.3.11 Borrowed Diagrams 201
6.3.12 Feedback 205
6.3.13 Tracking Visitors 206
6.3.14 ChangeLog 207
6.3.15 Search keyword(s) externally or internally 208
6.4 Credibility analysis of putative disease-causing genes using bioinformatics 209
6.4.1 Data collection 209
6.4.2 Pathogenicity analysis using bioinformatic tools 209
6.4.3 Data extraction from publications 230
6.4.4 Automated gene ranking 231
6.4.5 Validation of the method 244
6.4.6 User interface 247
6.4.7 Statistical methods 247
6.5 Development of a smartphone app for a genetics website, ALSoD 248
6.5.1 Searched online for direction 248
6.5.2 Choosing between a mobile website or an app 248
6.5.3 Optimization of webpage 248
6.5.4 Design heuristics 249
6.5.5 Mobile Device Detection 249
6.5.6 User Interface redirection 251
6.5.7 Compact version of website 251
6.5.8 Choosing to develop an app 251
6.5.9 Using WebView to open a website 252
6.5.10 Testing 252
6.5.11 App submission 254
6.5.12 Creating awareness 254
6.5.13 Feedback from app users 255
6.5.14 Analysis of visits 255

Chapter 7 RESULTS 256
7.1 Keeping up with genetic discoveries in amyotrophic lateral sclerosis: 256
7.1.1 Data Submission 256
7.1.2 Database Schema 258
7.1.3 Search terms on search engines 258
7.1.4 Contributors and Collaborators 260
7.1.5 Unique Identifiers 261
7.1.6 Graphical display of phenotypic data 261
7.1.7 Mapping variants on Google Earth 262
7.1.8 Meta-analysis and on-the-fly analysis of association data 263
7.2 ALSoD: A user-friendly online bioinformatics tool for Amyotrophic Lateral Sclerosis Genetics 264
7.2.1 Funding and Sponsorship 264
7.2.2 Programming 264
7.2.3 Web Design 264
7.2.4 On-the-fly Meta-Analysis of Genome Wide Association Studies in ALS 274
7.2.5 Collaborations and Embedded Tools 276
7.2.6 Integrated Bioinformatics Links 277
7.2.7 Feedback 278
7.2.8 Tracking Visitors 278
7.2.9 ChangeLog 281
7.2.10 Search keyword(s) externally or internally 281
7.3 Credibility analysis of putative disease-causing genes using bioinformatics 284
7.3.1 Data collection 284
7.3.2 Pathogenicity analysis using bioinformatics tools 284
7.3.3 Data extraction from publications 285
7.3.4 Automated gene ranking 285
7.3.5 Validation of the method 297
7.3.6 User interface 297
7.3.7 Statistical methods 297
7.4 Development of a smartphone app for a genetics website, ALSoD 298
7.4.1 Searched online for direction 298
7.4.2 Choosing between a mobile website or an app 298

7.4.3 Optimization of Webpages 298
7.4.4 Design Heuristics 299
7.4.5 Mobile Device Detection 299
7.4.6 Requesting Responses 299
7.4.7 App Development 299
7.4.8 Creating Awareness 299
7.4.9 Feedback From App Users 300
7.4.10 Analysis of Visits 300
7.4.11 Analysis to date (January 2014) 304
Chapter 8 DISCUSSION AND FUTURE WORK 307
8.1 Bioinformatics exploration of ALS genetics 307
8.2 ALSoD versus Others 308
8.2.1 Daily Updates 309
8.2.2 Data submission 310
8.2.3 Missing genetic data 310
8.2.4 Global representation 310
8.2.5 Popularity of database 310
8.3 Limitations of Genetic Databases 311
8.3.1 Continued Funding 311
8.3.2 Duplication of Efforts 311
8.3.3 Limited resources 311
8.3.4 Political boundaries on data sharing policies 311
8.4 Limitations of the ALS Online Database 311
8.4.1 Storage 311
8.4.2 Funding 312
8.4.3 Browser Incompatibility 312
8.4.4 Upgraded versions of third-party applications 312
8.5 Concluded Research Questions 313
8.6 Extra Development 314
8.6.1 Development of mobile computing applications 314
8.6.2 Expansion of phenotypic inclusion criteria to include frontotemporal dementia 314
8.6.3 Suggestions from feedback 314

8.7 Future of ALSoD 315
8.7.1 Collaborations 315
8.7.2 Larger Datasets 316
8.7.3 Graphical displays 316
8.7.4 Analysis 317
8.7.5 Development 317
Chapter 9 CONCLUSIONS 318
9.1 Summary for Chapter 3 318
9.2 Summary for Chapter 4 318
9.3 Summary for Chapter 5 320
9.4 Summary for Chapter 6 321
9.5 Conclusion 321
References 324
Appendix 1 – Data Submission Flowchart 346
Appendix 2 – Genome Wide Association Study Flowchart 347
Appendix 3 – Visitors' Statistics 348
Appendix 4 – Embedding Google Earth Flowchart 349
Appendix 5 – Flowchart for integrating Animal model into ALSoD 350
Appendix 6 - Data extraction from publications 351
Appendix 7 – Population Frequency script 359
Appendix 8 - Data extraction from publications 362
Appendix 9 – Population Frequency script using python 363
Appendix 10 - Process of installing and executing Eclipse application 365
Appendix 11 – SurveyMonkey screenshot 390
Appendix 12 – Script for Blog 391
Appendix 13 – Masterpage Websitemap 397
Appendix 14 – Tracking visitors script 408
Appendix 15 – Inserting association text files into SQL Server 412
Appendix 16 – ALTER TABLE for inserting ID and population column 414
Appendix 17 – Codes for creating scientific hyperlinks section 415
Appendix 18 - Claustalw and Jalview codes 431
Appendix 19 – GeneMANIA 435

Appendix 20 – Google Earth API 437
Appendix 21 – Python Script for Population Frequency 441
Appendix 22 – Perl Scripts to transpose, break, replace and join columns of a sequence 444
Appendix 23 – SQL Script for testing and verifying ranking system 456
Appendix 24 – Columns/Field list in each SQL table 458
Appendix 25 – T-SQL scripts in each SQL view 480
Appendix 26 – Disk Usage by Table in ALSoD 494
Appendix 27 – Label visibility on credibility analysis webpage 496
Appendix 28 – User interface ASP.NET code 498
Appendix 29 – MainActivity.Java for Android app 499
Appendix 30 – Possible graphs of 31 combinations for 5 populations 500
Appendix 31 – UCSC Genome Browser 504
Appendix 32 – UCSC Track for chromosome 9 554
Appendix 33 – Codes for searching SNP ids on GWAS page 558
Appendix 34 – Credibility analysis of genes on SPSS and spreadsheet 559
Appendix 35 – Populating Animal model table using T-SQL 560
Appendix 36 – GWAS method for 5 Populations 563
Appendix 37 – Updating database using T-SQL 573
Appendix 38 – Process of generating all possible mutations 627
Appendix 39 – Sample of letter sent to ALS Experts 644
Appendix 40 – Database Schema 645
Appendix 41 – List of Files 646
Appendix 42 - Recaptcha 661
Appendix 43 – Errors 664
Appendix 44 – Gantt Chart of Project 669
Appendix 45 – Glossary 670
Appendix 46 - Acronyms 675
Appendix 47 - Resources 676

Table of Figures

Figure 1: Mobile Marketing Statistics adapted from Netmarketshare webpage 64

Figure 2: Logging in section 110
Figure 3: Protected page after log in 111
Figure 4: Confirmation page 111
Figure 5: Mutation table on SQL Server 112
Figure 6: Gene submission page 1 112
Figure 7: Gene submission page 2 113
Figure 8: Gene submission page 3 114
Figure 9: Gene submission page 4 115
Figure 10: Gene submission page 5 116
Figure 11: Gene submission page 6 117
Figure 12: Gene submission page 7 118
Figure 13: Gene table on SQL server 119
Figure 14: Gene submission page on ALSoD 120
Figure 15: Patient table in SQL server 121
Figure 16: Patient submission page 1 122
Figure 17: Patient submission page 2 123
Figure 18: Patient submission page 3 124
Figure 19: Patient submission page 4 125
Figure 20: Patient submission page 5 126
Figure 21: Replicated mutation table on SQL server 127
Figure 22: Replicated mutation submission page 128
Figure 23: Database Schema 129
Figure 24: WGN-ALS logo 135
Figure 25: ENCALS logo 135
Figure 26: ALSA logo 136
Figure 27: MNDA logo 136
Figure 28: Clickable image using ImageMap control 136
Figure 29: Opening the Paint application 137
Figure 30: Using <asp:RectangleHotSpot> control 138
Figure 31: Full code for mapping image 138
Figure 32: Clickable image showing gene overview of MAPT 139
Figure 33: Using <asp:RectangleHotSpot> for logos 139

Figure 34: Calling Microsoft .NET framework 140
Figure 35: Microsoft SQL Server 2008 startup page 141
Figure 36: Microsoft SQL Server 2008 application 141
Figure 37: Using Python to map population frequency 142
Figure 38: WayBackMachine showing ALSoD homepage timeline 144
Figure 39: Home page 2007-2008 145
Figure 40: Home page 2009 146
Figure 41: Home page 2010 147
Figure 42: Home page 2011 148
Figure 43: Home page 2012 149
Figure 44: Home page 2013 150
Figure 45: Masterpage structure 151
Figure 46: Masterpage Header 151
Figure 47: Masterpage Footer 152
Figure 48: Masterpage Top section 152
Figure 49: Masterpage Left section 155
Figure 50: Chromosomal overview of chr 9 156
Figure 51: Flowchart for monitoring visitors 157
Figure 52: Analysis menu 162
Figure 53: Detailed Analysis submenu 162
Figure 54: Interaction submenu 163
Figure 55: Side by side comparison submenu 164
Figure 56: Pathogenicity prediction submenu 164
Figure 57: Credibility score analysis submenu 165
Figure 58: Summary menu 166
Figure 59: GWAS menu 167
Figure 60: Customizing UCSC tracks process 1 168
Figure 61: Customizing UCSC tracks process 2 169
Figure 62: Customizing UCSC tracks process 3 169
Figure 63: News menu 170
Figure 64: Data menu 171
Figure 65: Contributors menu 171

Figure 66: Feedback menu 172
Figure 67: GWAS flowchart 174
Figure 68: GWAS text file stored on SQL database 175
Figure 69: Genome Wide Association Study webpage 177
Figure 70: Whole Genome analysis 177
Figure 71: GWA_NOKEY table 178
Figure 72: Downloading java software 180
Figure 73: Haploview interface 181
Figure 74 : UCSC genome browser track configuration 185
Figure 75 : Flowchart of Google Earth 188
Figure 76: Navigating to Python application on desktop 192
Figure 77: Screenshot of executing python 192
Figure 78: Result of running Python script 193
Figure 79: Screenshot of reformatting Python output in spreadsheet 193
Figure 80: Accessing data from MGI website 195
Figure 81: Screenshot on the process of creating Animal model table 196
Figure 82: Chromosomal overview of all ALS-related genes 202
Figure 83: SOD1 subpsec score created in 2008 by Amy Butler 203
Figure 84: ALS-FTD relationship created in 2013 by Ashley Jones 204
Figure 85: Search externally and internally 208
Figure 86: Overview of credibility analysis method 209
Figure 87: Generating all possible gene mutation process 1 210
Figure 88: Generating all possible gene mutation process 2 211
Figure 89: Generating all possible gene mutation process 3 212
Figure 90: Generating all possible gene mutation process 4 213
Figure 91: Generating all possible gene mutation process 5 214
Figure 92: Pathogenicity prediction in PANTHER process 1 215
Figure 93: Pathogenicity prediction PANTHER process 2 216
Figure 94: Pathogenicity prediction in PANTHER process 3 217
Figure 95: Pathogenicity prediction in PANTHER process 4 217
Figure 96: Pathogenicity prediction in SIFT process 1 218
Figure 97: Pathogenicity prediction in SIFT process 2 219

Figure 98: Pathogenicity prediction in SIFT process 3 220
Figure 99: Pathogenicity prediction in SIFT process 4 221
Figure 100: Pathogenicity prediction in PolyPhen process 1 222
Figure 101: Pathogenicity prediction in PolyPhen process 2 223
Figure 102: Pathogenicity prediction in PolyPhen process 3 224
Figure 103: Pathogenicity prediction in PolyPhen process 4 225
Figure 104: Storing data from generated analyses into database tables process 1 226
Figure 105: Storing data from generated analyses into database tables process 2 227
Figure 106: Storing data from generated analyses into database tables process 3 228
Figure 107: Storing data from generated analyses into database tables process 4 229
Figure 108: Flowchart of pathogenicity prediction query 229
Figure 109: Criteria for Credibility score analysis 232
Figure 110: Number of affected patients in ALSoD 234
Figure 111: Number of mutations per gene 235
Figure 112: Number of cases recorded 236
Figure 113: Number of predicted pathogenic mutations by rank 237
Figure 114: Number of controls recorded 238
Figure 115: Number of mutations in same codon by rank 239
Figure 116: Number of patients with family history FALS 239
Figure 117: Number of ptients with SALS 240
Figure 118: Number of times mutation is replicated 241
Figure 119: Number of unique countries on genes 242
Figure 120: Survey displayed on webpage 245
Figure 121: Forced ranking output 246
Figure 122: Non-forced ranking 247
Figure 123: A method for redirecting to the appropriate view 250
Figure 124: Overview of mobile website development 250
Figure 125: Mobile view of website optimize 253
Figure 126: ALSoD Google play account showing number of downloads as at June 2013 254
Figure 127: Desktop view creating awareness using a Facebook and marquee 255
Figure 128: Submit new gene 256
Figure 129: Submit mutation data 257

Figure 130: Submit patient data 257
Figure 131: Submit replicated mutation data 258
Figure 132: Screenshot of search terms on pubmed 259
Figure 133: Screenshot of search terms on Google Scholar 259
Figure 134: Latest News webpage 260
Figure 135: Gene Overview page 262
Figure 136: Mapping variants on Google Earth 263
Figure 137: Sitemap of ALSoD 265
Figure 138: Interaction result using GeneMANIA 266
Figure 139: Side by side comparison result between FUS and TDP43 267
Figure 140: Detailed analysis of a selected gene 268
Figure 141: Summary report of genetic data 269
Figure 142: Displaying GWA result of 5 populations on Whole Genome 270
Figure 143: Displaying output of querying database for a single SNP 272
Figure 144: UCSC Genome Browser on chromosome 9 273
Figure 145: GWAS by Fogh 2013 273
Figure 146: Multiple Alignment on SOD1 277
Figure 147: Social media accounts of ALSoD 278
Figure 148: Visitors' statistics 279
Figure 149: Graph of Visitors' statistics 279
Figure 150: Google Analytics display 280
Figure 151: Search output using Google search engine interface for external search 282
Figure 152: Search output using Google search engine interface for internal search 283
Figure 153: Query result of predicted pathogenic FALS genes mutations 284
Figure 154: Number of affected patients 286
Figure 155: Number of mutations per gene 287
Figure 156: Number of cases recorded 288
Figure 157: Number of predicted pathogenic mutations by rank 289
Figure 158: Number of controls recorded 290
Figure 159: Number of mutations in the same codon by rank 291
Figure 160: Number of patients with family history 292
Figure 161: Number of patients with no family history 293

Figure 162: Number of times a mutation is replicated 294
Figure 163: Number of unique countries on genes....295
Figure 164: Number of publications 296
Figure 165: Reliability Analysis using Cronbach's alpha 298
Figure 166: Visits on ALSoD for 17 months on devices 304
Figure 167: First 10 Countries accessing database 305
Figure 168: First 10 Pageviews 306

Table of Tables

Table 1 : Origin of ALS-linked loci as at March 2014....26
Table 2: Non-locus specific mutation databases 54
Table 3: ALS-specific mutation databases....58
Table 4: Comparison between mobile website and app 65
Table 5: GWAS of 5 populations....181
Table 6: Comparison of website visits between the pre-mobile and post-mobile development. 300
Table 7: Referral traffic from search engines from August 2012 to June 2013 303
Table 8: First 10 Browser and OS used to access ALSoD from August 2012 to January 2014 305

Table of Equations

Equation 1 : Combination formula....179
Equation 2: Fisher's method 184

Document Information

Number of Pages: 680

Word count (with appendices and references): 112974

Word count (without appendices and references): 54066

File size: 71.7 MB

Acknowledgements

At long last, I can say “thank God for grace”. I appreciate the Almighty God for strength, sound health and the ability to end a long-awaited degree.

My profound gratitude goes to my unique Supervisor, Professor Ammar Al-Chalabi. I could not have done this without you nurturing and equipping me with all the necessary skills required for completing my programme despite your very busy schedule. Your fatherly support is incomparable and I sincerely appreciate all your efforts.

To my second supervisor, Dr John Powell, I am grateful for your immense support. You never turn me away even at your busy times. My impromptu turn-up at your office and your sincere comments, suggestions and praises have carried me through.

Thanks to Dr Nigel Leigh and Dr Radunivoc for bringing to life the ALS database idea in 1999. You gave me the raw materials to work on. To Dr Richard Wroe and Dr Amy Butler (who helped to proofread my thesis) who initially developed the website, you set the ball rolling for me and I cannot thank you enough.

I have been blessed with a wonderful team of colleagues: Dr Aleksy Shatunov, Dr Ashley Jones, Dr William Spatoveri, Dr Kuang Lin, Dr Conrad Iyiegbe, Dr Petra Petrouli, Dr Isabella Fogh, Michele Benigni and Simon Topp. Working with you all has exposed me to the benefits of working in a team. Thanks for all your help.

My family and friends have been very supportive during my period of study by asking about the progress of my programme and remembering me in prayers for he grace to carry on. Motivators worth mentioning are from my Aunts: Dr (Mrs) Adedayo, Dr (Mrs) Amao-Kehinde, Mrs Keshinro, Mrs Durojaiye, Ms Adesioye, Mrs Pat Babatunde and my uncle Mr Lekan Babatunde. Thank you very much and God bless you for your love. To TBBC family and UCU Alumni Family and Friends, I really appreciate your prayers and support.

To my late mum, Mrs A.B. Babatunde who experienced some of the times I had to write my M.Sc. exams while I was pregnant and helped to take care of her grand children - may your gentle soul rest in peace. I can never forget your love and care. To my dad, Mr E.O. Babatunde, who sacrificed his comfort to set my feet on the pathway to achieving greater height in life, I say a very big thank you. You shall reap the fruit of your labour over every of your children. My brother and sisters (Mr Babade Babatunde, Mrs Bisola Apata, Dr Bosede Omodara and Dr Bimbola Babatude), I love you all. My sisters-in-law and brothers-in-law (Mrs Ese Ndu, Mrs Folake Babade-Babatunde, Mr Ejiro Abel, Mr Edesiri Abel, Dr Niran Omodara, Mr Adewale Apata), you are a blessing to me. To my numerous cousins home and abroad (The Babatundes, The Durojaiyes, The

Adedayos and The Adesolas), I really appreciate you. To all members of True Believers Baptist Church UK, I sincerely appreciate your prayers and fellowship.

So many times I am asked how I am able to combine studentship with motherhood. My secret is my ever-loving, supportive, understanding, helpful and incomparable husband who gives me all the physical, emotional and spiritual support needed. What could I have done without you, my love, my partner, my backbone and soulmate. God bless and keep you. To my lovely damsels (Rebecca, Rachel, Ruth and Rhoda), mummy loves you very much. You are my crown. Thanks for being supportive in your own little ways.

Chapter 1 INTRODUCTION AND BACKGROUND

1.1 Health is wealth

"Health is wealth" is a common phrase often used in my growing-up days by friends and family especially when an action taken by an individual could affect his/her health but I never bothered searching for the origin of this statement until now. I only just realised that the exact quote is, "The greatest wealth is health" by a Roman Poet, Publius Vergilius Maro (from 70BC – 19BC commonly known as Virgil or Vergil) [1]. Every pregnant mother longs for a healthy baby at birth and every ill person yearns for a sound health. Any medical condition (either a disorder or disease) no matter how insignificant could rob an individual of his/her happiness. In the western world, studies have confirmed that population is growing older rapidly and a persistent drop in mortality is reported due to improved medical knowledge and services. Despite the breakthrough, increase in chronic diseases is experienced in many countries [2-4].

1.2 Diseases and Disorders

A disorder is a functional abnormality or disturbance while a disease is an abnormal condition affecting the body of an organism [5]. Searching through literature, I discovered there is a never-ending long list of diseases in the world today. Some are linked to genes and others are not, some are complex and others are not, some are curable and others are not, some are contagious and others are not. A chromosomal disorder in chromosome 21 is associated with patients with the Down's syndrome [6-8] while a non-genetic disease like Poliomyelitis threatened the health of children in the past but thanks to Dr Salk who created a vaccine to conquer Polio [9, 10]. A study conducted on chronic diseases selected 8 diseases for analysis –diabetes, heart disease, peripheral arterial disease, stroke, lung disease, joint diseases, back problems and cancer [3]. This is just a tip of the iceberg. Even OMIM which is a catalogue of Human genes and genetic disorders is only restricted to a list of diseases that are gene-related [11]. Geneticists have made significant advancement in discovering genetic root of diseases like Huntington's disease, Breast cancer, Alzheimer's disease while more research work is still being carried out on complex diseases like schizophrenia, bipolar disorder, and diabetes [12].

1.3 Motor Neuron Disease

Motor Neuron Disease (MND) is a neurological disorder affecting motor neurons, which control voluntary muscle movements. Forms of MND include Familial and Sporadic Amyotrophic Lateral Sclerosis (ALS), Spinal

Muscular Atrophy (SMA), Progressive Muscular Atrophy (PMA), Primary Lateral Sclerosis (PLS), Frontotemporal Dementia (FTD) and progressive bulbar palsy [13-15].

1.4 Amyotrophic Lateral Sclerosis

Amyotrophic Lateral Sclerosis (ALS) also known as Charcot disease [16] or Lou Gehrig's disease [17] or sometimes MND (Motor Neuron Disease) [18] , is a complex and fatal progressive neurodegenerative disease which kills patients within 2-5 years after diagnosis [19-22] and is a syndrome resulting in death from respiratory failure [18, 23]. ALS is known as Lou Gehrig's disease in the United Stated with association to a famous baseball player who died at the age of 37 years (2 years after he missed a game) due to this neurodegenerative disease [17, 24].

ALS is largely sporadic but with a minor proportion of familial cases which rarely exceeds 5% of all cases of ALS occur across population based studies[25]. ALS is the most common form of adult onset motor neuron degeneration, and affects adults in their mid-life [26, 27]. The causes of ALS are gradually being discovered but there is no known cure. Although there are other treatment strategies, the only effective drug treatment is riluzole, which helps prolong survival by about 3-6 months after 18 months treatment [28]. Non-invasive ventilation extends survival and improves quality of life [29].

1.5 Clinical representation of ALS

ALS affects motor nerve cells in the brain and the spinal cord resulting in muscle weakness and atrophy [16], leading to death through diaphragmatic paralysis. It involves both upper motor neurons (UMN) and lower motor neurons (LMN). UMN signs include hyperreflexia (overactive or over responsive reflexes), extensor plantar response (also known as Babinski response), increased muscle tone, and weakness. LMN signs include weakness, muscle wasting, hyporeflexia (below normal or absent reflexes), muscle cramps, and fasciculation (muscle twitch). Initial presentation varies. Affected individuals typically present with either asymmetric focal weakness of the extremities (stumbling or poor handgrip) or bulbar findings (dysarthria, dysphagia). Other findings may include muscle fasciculation, muscle cramps, and labile affect leading to involuntary emotional expression. Regardless of initial symptoms, atrophy and weakness eventually affect other muscles [30, 31].

1.6 Diagnosis

ALS is the most common form of adult onset motor neuron degeneration, and affects adults in their mid-life [26, 27]. Death from respiratory failure usually occurs within 3-5 years [19-22]. The occurrence and death rate of ALS have increased slightly over the past 50 years [32].

Because there is no diagnostic test, ALS poses a challenge to clinicians on how to diagnose the disease and the main guideline for the diagnosis and management of the disease is regularly updated [33, 34]. Even though early diagnosis is highly desired by clinicians, some diseases like Multiple Sclerosis, Pick's disease, Polymyositis, Myokymia etc can mimic ALS [33, 34]. In 1994, diagnosis criteria for ALS were proposed by the World Federation of Neurology. These are called the El Escorial criteria and were later revised in 2000 [35, 36]. Although largely helpful, the El Escorial criteria have been criticized for making recruitment to clinical trials difficult. In 2006, the Awaji algorithm was introduced to increase the eligibility of patients for clinical trials [37].

In recent times, the involvement of cognitive dysfunction in some ALS patients redirected researchers in finding a relationship between ALS and FrontoTemporal Dementia (FTD) (one form later localized to Chromosome 9) [38]. The diagnosis of FTD was based on the revised Neary criteria which were designed to improve clinical recognition of the disease [39, 40].

1.7 Prevention and Cure

"Prevention is better than cure" is an old adage commonly used and accepted as the truth but in the case of a complex disease like ALS, there is neither prevention nor a cure available. Studies are being carried out worldwide to determine ways of preventing ALS by observing possible patterns in the Age of onset of patients, duration of disease and site of onset in patient data which are curated into ALSoD.

There is no known cure available for the treatment of the disease but the use of Rilozule drug has some evidence of extending the survival of patients and improves their life quality [29, 41-43].

An announcement in November 2007 sparked a rush in treating ALS patients with lithium but was disproved in December 2008 from analysis directly carried out on patients [44-47]. Studies confirmed that Lithium did not demonstrate any valuable effect. It rather raised concerns on the reduced acceptability of lithium and its safety in ALS [48, 49].

1.8 Clinical Trials

A subcommittee of the MND Research Group of the World Federation of Neurology formed the Consortium on Clinical Trials in ALS in 1994 and then in 1998. The only approved treatment for ALS to date –as previously mentioned - which has gone through patient testing is Rilozule (*Rilutek*). Although this drug is not a cure, it extends the survival of patients. Further development is regularly carried out on potential drugs and compounds in various clinics across the globe for pharmacologic evaluation. Some of the drugs tested and/or being tested are Talampanel, Coenzyme Q10, Tamoxifen, Ceftriaxone, ONO-2506, AEOL 10150, Arimoclomol, Celastrol, Copaxone, IGF-1–viral delivery, Memantine, IGF-1 polypeptide, NAALADase inhibitors, scriptaid, Minocycline, Sodium Phenylbutyrate, Nimesulide, Thalidomide and Trehalose [50-54].

1.9 Palliative care / Equipment

The challenging nature of ALS makes it difficult for the physician to explain the diagnosis and prognosis of the disease to a patient thereby leaving patients uncertain about the diagnosis which often leads to more confusion when patients try to collect information from different sources [55, 56]. Palliative care as defined by the World Health Organization “is the active total care of patients whose disease is not responsive to curative treatment” which is aimed at “the best quality of life for patients and their families” [57]. Palliative care in ALS could increase survival if properly managed [55].

Quality of Life is “The extent to which hopes and ambitions are matched by experience”. Various equipment enhance the life quality of patients [33, 34, 58-60] which include percutaneous endoscopic gastrostomy (PEG), non-invasive ventilation (NIV), tracheostomy ventilation (TV), sniff nasal pressure (SNP), non-invasive positive-pressure ventilation (NIPPV) and invasive mechanical ventilation (IMV).

The ALS Functional Rating Scale (ALSFRS) was introduced as a rating instrument for measuring the degeneration of muscles in patients and this scale was later revised (ALSFRS-R) incorporating additional assessments and the necessity for respiratory support [61, 62].

1.10 Epistasis

The complexity of ALS genetics combines various challenging factors like epistasis making the disease complicated to study. The process by which genes interact to affect phenotypes and masking the presence of each other or combine to create a new attribute is known as Epistasis [63-65]. The occurrence of gene–gene interactions has been hypothesized to be at the foundation of many diseases common to humans even though current GWAS studies largely overlook its function [66, 67].

1.11 Epigenetics

Epigenetics is the process of modifying the DNA externally thereby affecting how cells interpret genes and not necessarily changing the DNA sequence. Because this study of heritable alterations in gene role occurs independently of changes to primary DNA sequence, it is proposed in a study that epigenetically silencing genes essential for the function of motor neuron could trigger Sporadic ALS. However, SOD1 and VEGF genes which were thought to be implicated in SALS do not have unusual methylation levels [68, 69]

1.12 Epidemiology

Epidemiology is the study of the pattern of a disease in different populations [70]. A study has pointed out that Italian Soccer professionals have a high risk of developing ALS [71-73]. A cluster of ALS patients who served in the military and veterans of the first Gulf war were also identified [74-76]. Together these evidences suggested that environment acting on genetic load is a likely critical contributor to ALS risk [77-80]. Examples of environmental factors affecting ALS include dietary, oxidative stress, occupation, head injuries, sports, pesticides, heavy metals and lifestyles [68, 69, 81].

1.13 Neuroimaging

Magnetic resonance imaging (MRI) uses strong magnetic fields and radio waves to produce detailed images of the inside of the body [82]. Most researchers support the use of MRI scan which it is thought could provide a substitute marker of ALS to help in detection of early stage disease and observe advancement and treatment reaction [83]. Studies have been carried out using neuroimaging to detect the characteristics of brain functions and structure by comparing the brain images of patients with ALS with healthy controls. It was discovered that the grey and white matter change in ALS and ALS-FTD patients [84-86].

1.14 In Vivo and in Vitro Models

There are more ALS publications in animal models and cellular models than in human models. The study of ALS toxicity in animal and cellular models aids in the investigation of pathological biomarkers for ALS. *Mus Musculus* (mice), rats (*Rattus norvegicus*), fruit flies (*Drosophila melanogaster*), zebrafish (*Danio rerio*), dog (*Canis familiaris*), pig (*Sus scrofa*), frog (*Xenopus Laevi*) etc. have been used mainly in SOD1, TDP-43 and FUS candidate gene studies [87, 88]. Tissues from cerebellum, spinal cord, brain tissues, blood etc have been examined to better understand the disease [89, 90].

1.15 Genetics of Amyotrophic Lateral Sclerosis

It is thought that the initial cases of ALS were defined in 1848 by Aran and in 1853 by Cruveilhier. But the first formal publication was written in 1874 describing a case of a female affected by an unknown disease in 1865 by the French physician, also regarded as the godfather of MND, Jean-Martin Charcot who lived from 1825 to 1893. Until the 1950s when the pedigrees of a family history was reported by Kurland and Mulder, ALS was not thought to have a familial component. A linkage study in 1991 suggested a familial ALS locus on chromosome 21, later identified to be the SOD1 gene. Eleven mutations were identified in 13 families [18, 23, 91].

About 5% of people with ALS have a family history of the disease. About 20% of such people harbour a mutation in the *SOD1* gene and a further 5% have mutations of *TARDBP*, and *FUS* [92-94]. Genes with strong but less robust evidence for being involved in familial ALS include *NEFH*, *ANG*, *OPTN* and *DAO* [95-98]. Genes for atypical forms of ALS, such as young onset, or slowly progressive ALS include *ALS2*, implicated in a recessive form of juvenile onset, mainly upper motor neuron disease, *SETX*, responsible for young onset, slowly progressive mainly lower motor neuron degeneration, and *VAPB*, responsible for various phenotypes of motor neuron disease, some similar to classical ALS, others largely lower motor neuron diseases.

Variants of all familial ALS genes are also found in sporadic ALS and account for about 10-15% of such cases. A recent meta-analysis of three twin studies has shown sporadic ALS has a heritability of about 0.61 (0.38, 0.78) confirming the validity of searching for genes associated with sporadic ALS[99]. A study suggests that the classification of genes as either sporadic or familial is artificial. Due to demographic drifts and family size, penetrance calculation leads to a false reclassification of a proportion of familial cases as sporadic [100].

There are two classical methods for the identification of such genes. The first is the candidate gene association study in which a gene considered to be relevant to ALS pathogenesis is tested for association of specific genetic variants with disease. The second is the genome-wide association study (GWAS) in which hundreds of thousands of gene variants are tested for association with ALS without prior hypothesis in a manner analogous to a genome-wide linkage study. Candidate gene studies have only yielded a few replicable genetic associations [101]. Examples of genes identified with this approach are *SOD1*, *TARDBP* (coding for TDP43), *FUS*, *ANG*, *NEFH*, *VEGF*, *ELP3*, *DPP6*, *ITPR2*, *UNC13A* and *ATXN2*.

Larger volumes of data are derived from rapidly increasing throughput from more individuals sequenced in a population with more sequenced genetic markers per individual with greater depth and accuracy [102]. Next-generation sequencing technology enables cheaply-produced massively sequenced data compared to the

conventional methods of sequencing [103] [104]. The detection of disease-specific biomarkers will make a way for the development of early diagnostic measures and substitute markers to monitor disease evolution and examine drug worthiness in clinical trials [105] [106]. It is becoming more difficult to keep abreast with this growing quantity of studies and enormous data regularly generated in ALS genetics research. A collaboration between ALS database teams working together to collate ALS results could help to keep researchers updated on current findings.

At least 15 fALS loci, under various modes of inheritance, have been identified by linkage studies, and pathogenic mutations have been described in 11 genes, SOD1, NEFH, ALS2, DCTN1, VAPB, SETX, ANG, TARDBP, FUS, OPTN and DAO, in fALS [107, 108].

Rare genes have a greater biological effect on patients but less impact on the disease due to the sample size whereas, common genes have a greater impact in terms of the numbers affected but with less biological effect on individual patients [109].

1.16 Environmental causes

Amyotrophic Lateral Sclerosis was observed in the Chamorro people of Guam with additional features similar to Parkinson's disease. The Chamorro tribe of Guam were afflicted with a neurodegenerative disease known as ALS-PDC. The major cause of this disease was associated with their diet of flying fox. Cycad seeds which are from neurotoxic plants are consumed by flying foxes causing ALS-PDC in the population. There is a biological amplification of the toxin in the fox tissues as they accumulate high levels so that people who eat them are then exposed to very high toxin levels.

Other suggested causes are chick pea and the custom of eating brains of dead ancestors. The β-N-Oxalylamino-l-alanine (BOAA) non-protein amino acid present in chickling peas [lathyrus sativus] causes neurolathyrism when eaten in excess and the toxins are concentrated in the brain tissues of patients [110, 111].

1.17 How credible are these reported genes?

The term 'credibility' was first used in a publication referring to how probable an association exists following a gathering of evidences. Pragmatic research and widely-agreed development are required to create a genetic model which combines biological and epidemiological evidences. The three Venice criteria used to determine how credible a gene is associated with a disease are: Amount of evidence (the larger the studies the better),

Replication (consistent replication across different populations) and Protection from bias (the less the prejudice in the studies the better) [112].

The discovery of many genes is helping researchers build a bigger picture of the pathogenic process involved and the connection between ALS and FTD helps to construct a better understanding for generating a treatment for ALS in the future [113]. What will be very useful in ALS studies is an analytical tool giving direction to new researchers in the field of ALS and encouraging current researchers to submit published and unpublished data for in-depth automated analysis of genetic data on a wider scope.

1.18 Origin of ALS-linked loci

There are over 20 ALS-linked loci, tens more linked to other motor neuron diseases (such as hereditary spastic paraparesis and distal spinal muscular atrophies) and several linked to related neurodegenerative diseases such as frontotemporal dementia which can affect the cognitive function of patients with ALS [14, 28, 113].

Table 1 : Origin of ALS-linked loci as at March 2014

Locus	Gene	Gene name	Chromosome	Mutations	Patients	Publications
ALS 1	SOD1	Cu/Zn superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))	21q22.11	177	374	125
ALS 2	ALS2	amyotrophic lateral sclerosis 2 (juvenile) homolog (human). Alsin	2q33.2	23	10	16
ALS 3	ALS3	Unknown	18q21	0	0	1
ALS 4	SETX	Senataxin	9q34.13	7	64	3
ALS 5	SPG11	spastic paraplegia 11 (autosomal recessive)	15q14	12	27	2
ALS 6	FUS	fusion (involved in t(12;16) in malignant	16p11.2	77	96	25

		liposarcoma)				
ALS 7	ALS7	Unknown	20p13	0	0	1
ALS 8	VAPB	Vesicle-associated membrane protein-associated protein B	20q13.33	2	19	6
ALS 9	ANG	Angiogenin	14q11.1	29	31	22
ALS 10	TARDBP	TAR DNA binding protein	1p36.22	50	91	29
ALS 11	FIG4	FIG4 homolog, SAC1 lipid phosphatase domain containing (S. cerevisiae)	6q21	10	9	1
ALS 12	OPTN	optineurin	10p13	37	19	9
ALS 13	ATXN2	ataxin 2	12q23-q24.1	1	10	3
ALS 14	VCP	valosin-containing protein	9p13	7	9	2
ALS 15	UBQLN2	ubiquilin 2	Xp11.21	6	40	4
ALS 16	SIGMAR1	sigma non-opioid intracellular receptor 1	9p13	1	0	1
ALS 17	CHMP2B	chromatin modifying protein	3p11.2	2	0	3

		2B				
ALS 18	PFN1	profilin 1	17p13.3	4	5	7
ALS 19	ERBB4	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 4	2q33.3-q34	2	6	1
ALS 20	HNRNPA1	heterogeneous nuclear ribonucleoprotein A1	12q13.1	1	0	1
ALS-FTD 1	ALS-FTD1	Unknown	9q21-q22	0	0	1
ALS-FTD 2	C9orf72	chromosome 9 open reading frame 72	9p21.2	1	76	70
ALS	UNC13A	unc-13 homolog A (C. elegans)	19p13.12	0	0	6
ALS	DAO	D-amino-acid oxidase	12q24	2	2	2
ALS	DCTN1	Dynactin	2p13	6	6	4
ALS	NEFH	neurofilament, heavy polypeptide 200kDa, heavy chain	22q12.1-q13.1	11	10	10
ALS	PRPH	peripherin	12q12	0	0	2
ALS	SQSTM1	sequestosome 1	5q35	16	14	2
ALS	TAF15	TAF15 RNA polymerase II, TATA box binding protein	17q11.1-q11.2	7	4	1

		(TBP)-associated factor, 68kDa				
ALS	SPAST	Spastin	2p24	0	0	3
ALS	ELP3	elongation protein 3 homolog (S. cerevisiae)	8p21.1	0	0	1
ALS	SPAST	Spastin	2p24	0	0	3
ALS	ELP3	elongation protein 3 homolog (S. cerevisiae)	8p21.1	0	0	1

1.18.1.1 ALS1/SOD1 [147450]

Cu/Zn superoxide dismutase gene is associated with autosomal dominant form having 23 exons on chromosome 21q22.1 consisting of 154 amino acids with 160 mutations reported so far.

In 1991, Siddique mapped Autosomal dominant forms to chromosome21q21 [114] .

In 1993, Rosen described eleven disease-associated mutations (Gly37Arg, Leu38Val, Gly41Ser, Gly41Asp, His43Arg, Gly85Arg, Gly93Cys, Gly93Ala, Glu100Gly, Leu106Val, Ile113Thr) in 13 different FALS families [91]. Aoki and Ogasawara found an exon 2 novel point mutation His46Arg in two unrelated Japanese families [115]. Deng detected an exon 1 mutation Ala4Val and exon 5 mutations (Ile144Phe, Val148Gly) [116].

In 1994, Aoki found a novel H46R mutation in 2 FALS Japanese families [117]. Elshafey identified a new missense point mutation (Gly93Arg) in exon 4 in a family [118]. Esteban found two novel mutations in exon 4 (Gly93Asp, Ile112Thr) and a new intron 3 polymorphism (34basepair A to C) in FALS patients but not in 169 sporadic ALS patients nor in 100 normal controls [119]. Hirano identified a heterozygous mutation (Val7Glu) in 2 members of a Japanese family[120]. Jones reported a new missense mutation (Asp101Asn) in exon 4 of a sporadic 53 year old ALS patient of Asian origin after screening a cohort of 46 sporadic and 10 familial ALS patients [121]. Jones again detected a novel Glu21Lys mutation and a previously detected Ile113Thr in three

other patients [122]. Kostrzewa recognized an exon 4 mutation (Arg115Gly) in a FALS family which was not detected in 90 controls nor in 212 normal individuals [123]. Nakano discovered an exon 1 novel missense mutation (Ala4Thr) in 2 FALS patients from one Japanese family after studying 17 SALS patients and 9 FALS patients [124]. Pramatarova found a two base pair deletion (L126delTT) in a FALS patient [125]. Rainero discovered a Gly41Ser mutation in exon 2 from 8 patients in an Italian pedigree comprising 75 members distributed in five generations [126]. Siddique identified 14 previously found mutations (A4V, A4T, G37R, G41D, H43R, G85R, G93A, E100G, L106V, I113T, L144F, V148G) and two new mutations, V14M and L84V. Mutation A4V is the most frequent one which occurred in 14 out of 29 FALS families [127]. Suthers reported a mutation Ile113Thr [128]. Takahashi also identified a previously reported A4T mutation in an autopsied patient with FALS [129]. Tsuda investigated the effects of a Gly41Ser mutation found in a FALS pedigree [130].

In 1995, Andersen found an exon 4 homozygous Asp90Ala mutation in 14 patients (from 4 unrelated ALS families and 4 sporadic ALS patients) from Sweden and Finland [131]. Aoki reported a new point mutation Leu84Val in exon 4 in a Japanese FALS patient but not observed in the 57 normal Japanese control subjects [132]. Deng identified 15 different mutations (Ala4Val, Ala4Thr, Val14Met, Gly37Arg, Gly41Asp, His43Arg, His46Arg, Leu84Val, Gly85Arg, Gly93Ala, Glu100Gly, Leu106Val, Ile113Thr, Leu144Phe, Val148Gly) in 37 FALS families out of a total of 222 families screened [133]. Enayat detected 2 novel mutations (in exon 2-His48Glu, exon 5- Asp125His) and four formerly detected mutations (exon 4-Ile112Thr, Ile113Thr, Gly93Arg, exon 5-Ile149Thr) in UK families [134]. Ikeda identified a novel Val148Ile mutation in exon 5 [135]. Ikeda again detected an exon 4 novel Ile104Phe mutation in affected members of three generations of Japanese kindred [136]. Orrell reported Gly93Arg mutation and Ile133Thr mutation in another publication [137, 138]. Pramatarova examined 114 unrelated affected individuals and 67 controls identifying previously reported mutations (Ala4Val, Gly37Arg, Gly93Cys, Ile113Thr, Leu144Phe) and two novel mutations (Ile149Thr, Asn139Lys) in a total of 15 families [139]. Sapp found 3 mutations (Leu144Ser, Ala145Thr and intron 4 T-G 10bp before exon 5) [140]. Sjalander identified a heterozygous and homozygous missense mutation Asp90Ala (D90A) in exon 4 from Swedish and Finnish populations [141]. Yulug detected E100G, I113T and a new mutation Asp101Gly [142].

In 1996, Abe reported 4 missense mutations (His46Arg, Leu84Val, Ile104Phe, Val148Ile) in 5 Japanese families [143]. Andersen described 36 patients (6 SALS cases and 30 FALS cases from nine families) characterized by a distinct phenotype associated with homozygosity for an Asp90Ala mutation [144]. Hosler identified 3 novel mutations (Gly93Val, Asp124Val, Glu133) and 2 variants (Ser59Ser, Ala140Ala) [145]. Ince described an autopsy case with the mutation E100G (exon 4) [146]. Kostrzewa revealed an exon 5 mutation

Ile151Thr in FALS [147]. Morita discovered a novel two-base (TGC to TTT) mutation Cys6Phe in a Japanese family [148]. Orrell examined a sporadic patient originally from Pakistan with a previously identified Asp101Asn mutation [149]. Siddique discovered the E100K mutation [150].

In 1997, Andersen found 5 different mutations (Ala4Val-Central Sweden, Val14Gly-Central Sweden, Asp76Tyr-Denmark, homozygous Asp90Ala-Finland and Sweden, heterozygous Asp90ala-Northern Finland and Gly127insTGGG-Denmark) [151]. Bereznai found a novel mutation L8Q in an Austrian family [152]. Cudkowicz screened 290 families detecting 18 mutations (Ala4Val, Ile113Thr, Glu100Gly, Leu38Val, Gly93Ala, Gly37Arg, Gly41Asp, Gly93Cys, Gly41Ser, His43Arg, Leu106Val, Val148Gly, Gly93Asp, Ile112Thr, Leu144Ser, Ala145Thr, Intron Phe-Leu-Gln, Glu133) in 68 families [153]. Jackson analyzed 155 cases identifying previously reported mutations (Asp90Ala, Ile113Thr) and a novel mutation (Val118Lys) in exon 4 in a single SALS patient [154]. Kawamata found a novel Gly16Ser mutation in a SALS patient with young onset [155]. Kawata identified a missense mutation (Gly93Ser) in exon 4 of a 48-year-old Japanese man with familial amyotrophic lateral sclerosis (FALS) [156]. Kikugawa performed mutational analyses of the SOD1 gene of 23 patients (three familial cases and 20 sporadic cases) with ALS from the Kii Peninsula and its vicinity and identified the same missense mutation (Ile113Thr) in exon 4 as a heterozygous state [157]. Maeda found mutation N86S in a FALS patient [158]. Orrell discovered a novel Gly108Val mutation identified in a family manifesting ALS in 3 generations [159]. He also found a novel heterozygous mutation in exon 3 (Gly72Ser) in a family with 2 individuals affected by ALS [159]. Orrell again investigated 71 families identifying 10 previously reported mutations(His48Gln, Gly93Arg, Gly93Val, Glu100Gly, Asp101Gly, Asp101Asn, Gl108Val, Ile113Thr, Asp125His, Ile149Thr) and a novel insertion (132insTT)[160]. Shaw P. reported E100G (in 1 FALS case), I113T (in 1 FALS and 1 SALS), and an exon 5 in a SALS (3'UTR 816-819 deletion) [161]. Shaw C. detected appoint mutation His48Glu in a patient with rapidly progressive FALS [162].Watanabe discovered a novel missense point mutation S134N in a FALS patient [163]. Zu found 2 new mutations L126Z and intronic mutation from A-G in intron 4, 11 bp has been opened by removing the loop forming one entire side upstream in 2 unrelated patients [164].

In 1998, Boukaftane screened affected individuals from 70 unrelated FALS kindreds and identified 10 mutations, 6 of which are novel (Glu21Gly, Leu38Arg, Glu49Lys, Leu84Phe, Leu67Arg, intron4del30bp) [165]. Hayward detected a homozygous Asn86Ser mutation in a juvenile onset FALS case [166]. Morita found a novel Asp90Val mutation in a Japanese family with ALS [167]. Nakanishi obtained blood samples from FALS cases detecting G37R, A4S, H46R, I149T [168]. Shaw discovered 8 of 38 patients with FALS and 5 of 175

patients with SALS had missense mutations with 2 novel mutations (L84F, G72S) identified [169]. Cudkowicz examined 11 subjects with A4V mutation as a replicated study [170].

In 1999, Aguirre studied 23 patients from 11 ALS families and 69 SALS patients of Belgian origin identifying 3 different mutations (L38V, D90A and G93C) in seven families [171]. Ceroni examined 13 family members and 6 of 10 unrelated patients with sporadic cases in Italy and discovered that all familial ALS patients and one of the six sporadic patients carry the same L84F missense point mutation in exon 4 of the SOD1 gene [172]. Kohno identified a novel missense mutation (Cys6Gly) in exon 1 in a Japanese woman and her family [173]. Orrell studied 17 families from the UK identifying 12 different mutations (H48Q, G72S, G93R, G93V, E100G, D101N, D101G, G108V, I113T, D125H, I149T – and also an insertion mutation – 132insTT) [174]. Penco reported a new missense mutation (Gly12Arg) in exon 1 in a 67-year-old patient with familial ALS (FALS) [175].

In 2000, Cervenakova reported that a D101N mutation in exon 4 of the SOD1 gene was identified in a PMA/ALS patient and in one of her three unaffected children [176]. Shimizu found a novel V118L mutation in a sporadic ALS patient [177]. Slominsky discovered a missense point mutation Asp90Ala in two patients and a novel intronic point mutation IVS3 + 35A>C in two typical sporadic ALS patients [178]. Weber studied 8 patients with homozygous D90A mutation, compared the results with those examined from 12 SALS and 11 healthy controls [179]. Winter found the D90A mutation in Germany [180].

In 2001, Gellera found mutations in 5 of 34 unrelated FALS and in 2 of 44 SALS patients in an Italian population with 2 FALS cases having A4V, 1 FALS case of with L84F. One FALS patient carried new G12R substitution in exon 1, one FALS patient with a F45C substitution in exon 2, 1 SALS patients carried the homozygous D90A, 1 SALS patient had the heterozygous I113T mutation and 1 SALS patient with an A95T amino acid substitution [181]. Hand described a French amyotrophic lateral sclerosis (ALS) family with two distinct mutations (heterozygous for the D90A mutation and also carry a single copy of a novel SOD1 mutation, D96N) [182]. Kato identified a homozygous Leu126Ser mutation in a Japanese family [183]. Mase investigated a large family of Istro-Rumanian origin in 18 cases throughout 6 generations discovering a Leu144Phe mutation [184]. Murakami identified a novel missense mutation Leu126Ser [185]. Takehisa also claimed to have discovered 2 Japanese families with a novel mutation TTG to TCG missense heterozygous mutation in the exon 5 causing substitution of leucine for serine at codon 126 (Leu126Ser) [186].

In 2002, Alexander discovered an H80A mutation in a patient with genetically proven sporadic ALS [187]. Garcia-Redondo found G37R in exon 2, N65S in exon 3 and I112M in exon 4 in Spanish patients with

sporadic or familial ALS [188]. Iwai examined the G93S mutation in Japanese population [189]. Jonsson analyzed D90A heterozygotes from recessive and dominant ALS pedigrees [190]. Naini reported a familial amyotrophic lateral sclerosis (FALS) patient presenting a late-onset disease associated with a missense mutation (Ala140Gly) in exon 5 [191]. Segovia-Silvestre found a D76V amino acid change on a large family with 15 affected individuals spanning five generations [192].

In 2003, Aksoy reported an exon 1 mutation A4T in a large kindred of FALS [193] while Andersen found 16 novel exonic mutations (L8V, F20C, Q22L, H48R, T54R,S59I, V87A, T88DTAD, A89T, V97M, S105delSL,V118L, D124G, G141X, G147R, I151S) [194]. Ferrera examined the first Italian FALS comprising of 7 affected members in 6 generation pedigree with a Leu144Phe mutation [195]. Kim observed a novel mutation Gly10Val in a Korean family [196]. Mayeux discovered a new mutation N19S in 2 SALS patients after examining 331 SALS cases though it was also found in 1 healthy control [197]. Rezania presented three members of a pedigree with familial amyotrophic lateral sclerosis (FALS) who have a rare mutation in exon 4 of Ala89Val [198].

In 2004, Niemann found R115G mutation in 4 FALS patients in a German population [199]. Nogales-Gadea discovered an exon 5 mutation N139H in a Spanish family [200]. Ohi analyzed a Japanese family carrying the His46Arg mutation [201]. Sato identified 2 novel mutations (Asp101His and Gly141Glu in exons 4 and 5 respectively) [202]. Shi identified an Insertion mutation of exon 2 in an ALS family in Chongqing, China [203]. Tan described a patient with FALS having a novel missense mutation (Asp101Tyr) in exon 4 [204].

In 2005, Antonyuk examined the His46Arg mutation in familial ALS [205]. Battistini studied 264 Italian patients (with 225 SALS and 39 FALS) identifying G12R, G41S, L114F, D90A, S59S, IVS3: + 34 A-->C, IVS3 + 62 T-->C mutations [206]. Sato examined 29 patients with ALS and 4 carriers having previously detected Ala4Ser, Ala4Thr, Ala4Val, Val14Gly, Gly37Arg, Gly41Asp, His46Arg, Asp76Tyr, Asn86Ser, Ala89Val, Asp101His, Ser105Leu, Ile113Thr, Ile113Phe, Gly114Ala, Arg115Gly, Leu1262bp del, Leu126Ser, Gly127insTGGG, Ser134Asn, Gly141Glu, Leu144Phe, Ala4Val, Asp76Tyr, Ile113Thr [207]. Zhang found a novel mutation Glu133Val in FALS patients in China [208].

In 2006, Andersen discovered novel N139D, G147D mutations amongst other previously detected SOD1 mutations [209]. Corrado found a new mutation K136X and replicated mutations N65S and A95T in Italian patients [210]. Fong reported a large multigenerational Chinese FALS kindred with I149T mutation [211]. Gamez carried out a mutational analysis and identified 5 out of 30 FALS and 4 out of 94 SALS patients with Gly37Arg (G37R), Asp76Val (D76V), Ser105Leu (S105L), Ile112Met (I112M) and Asn139His (N139H)

mutations [212]. Li Xiaoguang found a heterozygous mutation, GAA to GTA, causing the substitution of valine for glutamic acid at codon 133 (Glu133Val) in exon 5 in five out of 16 individuals from one Chinese family [213]. Regal conducted further research on 20 patients with G93C mutation [214]. Stewart analyzed a 71 year old woman revealing a novel mutation G72C [215].

In 2007, Beck found a new heterozygous N86K mutation in a patient with an inherited ALS [216]. Holmoy reported the first case of FALS caused by H46R in a patient of Pakistani descent [217]. Kim H.Y. identified Phe20Cys (F20C) mutation in a Korean family [218]. Kim W. revealed Val10Gly, IVS4+15_16insA, IVS4+42delG and IVS4+59_60insT variants [219]. Naini identified a novel Asp109Tyr mutation in exon 4 (D109Y) in a sporadic ALS woman [220].

In 2008, Broom examined the homozygous 50bpdel mutation [221]. Cao studied the structure of G85R variant in familial ALS [221]. Corti found a new heterozygous mutation Q22R in codon 22 [222]. Eisen discovered 6 missense mutations (A4V, G72C, D76Y, D90A, C111Y, I113T) in 13 patients of both FALS and SALS cases [223]. Suzuki described a 39 year old Japanese woman with FALS identifying a missense exon 4 mutation (Gly93Ser) [224].Takahashi conducted an analysis of 35 patients with sporadic ALS and 238 controls revealing a previously known SOD1 mutation, S134N [225].

In 2009, Syriani identified E22G sequence change in 3 affected ALS patients [226]. Winkler studied 2 mutations His46Arg (H46R) and His48Gln (H48Q) claimed to be pathogenic [227]. Zinman found a homozygous novel 6bp deletion in exon 2 (delG27/P28) after examining 12 members of a Canadian ALS family of Filipino origin [228].

In 2010, Battistini found a severe familial ALS with a novel exon 4 mutation (L106F) [229]. Felbecker reported four ALS pedigrees from Finland, France, Germany and Sweden with either the D90A or E100K SOD1 mutations [230]. Georgoulopoulou found a new mutation D11Y in a sporadic ALS patient [231]. Giannini first described two, apparently sporadic, Italian ALS patients heterozygous for the D90A mutation [232]. Holmøy described a SALS patient with a novel G127R mutation at position 382 [233]. Kobayashi discovered a Japanese FALS patient with Gly72Ser mutation in SOD1 gene [234]. Laaksovirta analysed 405 ALS patients with 497 controls from a sample of 442 ALS patients and 521 controls after filtering data for quality control; and discovered rs13048019 with a p-value of 2·58×10(-8) in chromosome 21q22 [235]. Millecamps screened a cohort of 162 families enrolled in France and 500 controls identifying 31 pathogenic missense mutations found in 36 patients (20 SOD1 - E21G, G37R, L38R, G41S (2 patients), H46D, P66R, D83G, L84F, L84V, N86S, G93C, G93A, G93D, G93V, V118L, N139D, G147D (2 patients) and I151T , 1 ANG, 1 VAPB, 7 TARDBP and

7 FUS) [236]. Origone found a novel Gly147Ser mutation in an Italian sporadic ALS [237]. Rabe analyzed 217 families from Germany (n = 213), Austria (n = 2) and Switzerland (n = 2) discovering previously detected mutations Q22L, L38V, H46R, L84F, N86D, D90A, G93A, E100K, I104F, I113T, R115G, L144F, V148G, I149T and I151T [238]. Ricci reported a novel exon 1 G10R mutation in an examined 68 year old female with FALS [239]. Synofzik presented a G93A mutation from a large German pedigree [240]. Takazawa found a new G85S mutation in exon 4 [241]. Tsai screened a cohort of 15 index patients of Taiwanese Han Chinese descent with adult-onset FALS identifying 7 different mutations in 8 patients, including 3 in SOD1 (G85R, T137R, and G138E), 2 in exon 15 of FUS and 2 in exon 6 of TARDBP [242]. Van Es screened a total of 451 sporadic and 55 familial patients detecting a novel mutation I99V, a homozygous D90A mutation identified in SALS and one FALS patient had a heterozygous D90A mutation [28]. Visani found a novel variant T137A mutation in two members of an Italian family [243].

In 2011, Baek found a novel A4F mutation in FALS in 6 family members [244]. Brotherton found Cys6Ser mutation [245]. Conforti reported a G61R mutation in a sporadic ALS patient [246]. Del Grande described a novel L67P mutation located in exon 3 [247]. Hermann reported a heterozygous I113F mutation in FALS [248]. Vela examined the N19S mutation in a family with 4 ALS patients [249]. Visani found T137A mutation in an Italian family[243]. Zhao identified two previously reported mutations Cys6Phe and Leu84Phe in two Chinese families [250].

In 2012, Berdynski replicated a study of G41S mutation in a large Polish family [251]. Brown found Val95Ala*, His121Leu, Asp126ThrfsX24, Gly142Ala amongst other replicated variants previously found [252]. Hu identified a novel frameshift mutation in a Chinese family on positions 29-35 converting Val-Lys-Val-Trp-Gly-Ser-Ile to Asp-Glu-Gly-Val-Gly-Lys-His [253]. Keckarević found a novel P66S mutation in exon 3 [254]. Origone analyzed the novel Gly147Ser mutation in a sporadic ALS Italian patient. [255]. Piaceri found a single SOD1 mutation (Asp109Tyr) and 3 TARDBP mutations in 5 SALS patients from 61 Italian ALS patients [256].

In 2013, Diard-Detoeuf found a novel heterozygous I18del variant [257]. Dangoumau discovered a novel heterozygous V31A variant in a Bulgarian man [258]. Nakamura conducted an autopsy on a SALS case with the replicated I113T mutation previously discovered [259]. Naruse studied a Japanese family reporting a novel A4D mutation [260]. Klein identified a novel Phe64Leu mutation [261] while Wang found a novel C7W mutation in a FALS pedigree [262]. Cui reported a new mutation S134T in a female ALS Chinese patient [263]. Trojsi found in an Italian kindred the replicated Ile149Thr mutation [264]. Kuzma-Kozakiewicz found a recurrent K3E mutation found in Polish and Japanese ALS patients [265].

1.18.1.2 ALS2

Alsin gene is a causative gene for a juvenile autosomal recessive form of motor neuron diseases (MNDs) having 33 introns and 34 exons on chromosome 2q33.2 consisting of 1657 amino acids with 23 mutations reported so far. Losing ALS2 function results in numerous recessive motor neuron diseases [266].

Refsum, Ben Hamida and Gragg originally described the ALS2 gene when patients with hereditary motor system diseases were studied [267-270].

In 1992 and 1994, Hentati applied linkage analysis to a large Tunisian family suggesting that the ALS2 locus lies in the 8cM segment flanked by D2S155 and D2S115 [271, 272].

In 1998, Hosler confirmed the previously reported linkage of the chromosome 2q33-35 region while analyzing additional markers [273].

In 2001, Hadano and Yang first linked the recessive form of juvenile ALS type 3 to chromosome 2q33-35 and subsequently also linked juvenile-onset PLS (JPLS) to 2q33 reporting two deletion mutations (A46fsx50 which is 138delA and L623fs×646 which is 1867-1868delCT) in two Saudi families (North African and Middle Eastern origin) [274, 275].

In 2002, Eymard-Pierre reported that Alsin mutations are responsible for a primitive, retrograde degeneration of the upper motor neurons of the pyramidal tracts, leading to a clinical continuum from infantile (IAHSP) to juvenile forms with (ALS2) or without (JPLS) lower motor-neuron involvement. Four alsin mutations (3742delA, 1471delGTTTCCCCCA, 2660del AT, 1130delAT) were discovered in four families that lead to abnormal short and long ALS2 transcripts and truncated alsin proteins suggesting that IAHSP is allelic to juvenile ALS and JPLS [276].

In 2003, Al-Chalabi's results showed that variants of the ALS2 gene are not a common cause of a predominantly young onset, upper motor neuron disease phenotype of sporadic ALS, nor are they associated with a more-typical phenotype. Twenty-one variants including 5 exonic,13 intronic and 3 non-SNPs (ivs1–57(T)11–13 , ivs2+62 t>c , ivs3–10(T)9–10 , 1102g>a V368M , ivs6+49 c>g , ivs10–62 c>t, ivs11–39 c>a, 2466g>a V822V, ivs13–142 t>a, ivs13–132 t>c, 2796c>t S932S, ivs16–30_–26del, ivs19+40 c>t, ivs20–75 a>g, 3885g>a A1295A, ivs25–76 t>c, 4015c>t L1339L, ivs26–64 g>a, ivs29+7 g>a, ivs30–69 t>a, ivs34+29 t>c) were found in 12 individuals [277]. Devon has revealed that ALS2 mutations are much more common than previously supposed due to the subsequent discovery of further ALS2 mutations within multiple different populations and a novel homozygous C3115T variant (R998X) was discovered in a Jewish family [278]. Gros-

Louis reported the identification of a novel ALS2 mutation in a large consanguineous Pakistani kindred with infantile onset autosomal recessive complicated HSP [279]. Hand screened 95 unrelated familial, 95 unrelated sporadic, and 11 early-onset ALS patients detecting 23 novel sequence variants; however, none is disease-associated [280]. Nagano screened 3 Japanese patients finding no deletion mutation suggesting that deletion mutations in ALS2 gene detected in ALS2 patients seem to be uncommon in Japanese patients [281].

In 2005, Kress presented an ALS2 Turkish male patient with a novel homozygous deletion in exon 4 identified as 553delA [282].

In 2006, Panzeri described the first homozygous missense which falls within the RCC1 domain in ALS2, c.1619 G >A which is p.G540E mutation identified in a patient [283]. Eymard-Pierre reported a novel ALS2 missense homozygous G669A mutation in patients from a consanguineous Turkish family affected by IAHSP, and demonstrated the instability of this ALS2 mutant protein [284].

In 2008, Verschuuren-Bemelmans recognized 10 homozygous variants. Nine were previously described and a nonsense mutation c.2143C4T in exon 10 of the ALS2 gene in two siblings with infantile-onset ascending spastic paralysis which predicted chain termination at amino-acid position 715 of the gene product ALSIN (p.Gln715X) was discovered [285].

In 2009, Shirakawa also reported heterozygous ALS2 mutations (V1189WfsX19, G1172EfsX29) in Japanese siblings [286].

In 2011, Mintchev reported a new homozygous splice-site frameshift mutation c.2980-2A>G (p.993fsX7) on intron 17 thereby causing a loss of exon 18 in a Cyprus family [287].

1.18.1.3 ALS4 - SETX

Senataxin gene is associated with autosomal dominant juvenile form of ALS having 24 exons on chromosome 9q34.13 consisting of 2677 amino acids with 7 mutations reported so far. It was suggested that juvenile-onset sporadic ALS and hereditary motor neuropathy patients with syndromes of ataxia and motor neuron disease should be examined for SETX mutations [288].

In 1998, Chance first linked ALS4 locus to markers from chromosome 9q34 after studying 11-generation pedigree with autosomal dominant, juvenile-onset motor-systems disease [289].

In 2004, Chen detected 3 missense mutations (L389S, T3I, R1236H) in a Maryland and Belgian family when 19 genes were tested within the ALS4 interval on 9q34 [290].

In 2009, Zhao screened 45 SALS and 200 controls identifying a novel variation Thr118Ile in a 42 year old SALS Chinese patient [291].

In 2011, Hirano detected c.4660T > G transversion mutations in two relatives of a patient [288].

1.18.1.4 ALS5 - SPG11

Spatacsin gene is associated with autosomal recessive juvenile ALS having 40 exons on chromosome 15q14 consisting of 2330 amino acids with 12 mutations reported so far.

In 2010, Orlacchio analysed patients of 25 unrelated pedigrees identifying 12 mutations (Gln40X, Trp89X, Met245Valfs, Arg1992X, Leu733X, Ile870Valfs, Arg1026fs, Val1468Leufs, IVS30 + 1 G > A, Tyr1990X, Val2053Met, Val2344Cysfs) [292].

In 2012, Daoud discovered a novel deletion variant, c.5199delA in exon 30 using exome sequencing [293]

1.18.1.5 ALS6 - FUS

Fusion gene is associated with autosomal dominant form having 15 exons on chromosome 16p11.2 consisting of 347 amino acids with 65 mutations reported so far.

In 2003, Abalkhail identified a putative locus on chromosome 16q12.1-q12.2 after carrying out a preliminary genome screen, using a U.K. resource of families lacking SOD1 mutations [294]. Ruddy identified two large European families with ALS disease were demonstrated only across regions of chromosome 16 [295]. Sapp performed a genetic linkage screen in 16 pedigrees with FALS and identified novel ALS loci on chromosomes 16 and 20 [296].

In 2009, Belzil screened the entire gene in a cohort of 200 patients with ALS identifying 3 different mutations (S57delTCT, R521C, R521H) in 4 different patients, including 1 3-bp deletion in exon 3 of a patient with sporadic ALS and 2 missense mutations in exon 15 of 1 patient with familial ALS and 2 patients with sporadic ALS [297]. Chio identified a heterozygous p.R514S missense mutation in a family of northern Italian origin, and a heterozygous p.P525L missense mutation in a family of Sicilian origin [298]. Corrado examined 1009 patients (45 FALS and 964 SALS) with another 293 of the SALS patients identifying seven missense mutations (G191S, R216C, G225V, G230C, R234C, G507D and R521C) in nine patients (seven SALS and two FALS), and none in 500 healthy Italian controls [299]. Damme sequenced the FUS gene in a cohort of patients from Belgium with familial ALS and identified a R521H mutation in 4 patients [300]. Kwaiatkowski sequenced all 15 exons in 81 other unrelated FALS cases and 293 sporadic ALS (SALS) DNA samples, and an additional 209 ALS families were screened for mutations in exon 15 detecting 13 different mutations

(H517Q, R521G, insGG, delGG, R244C, R514S, G515C, R518K, R521C, R521H, R522G, R524T, R524S, P525L) in 17 different FALS families [301]. Ticozzi identified 4 distinct missense FUS gene mutations (R521C, R521G, G156E, R234L, R522R) and 5 novel variants in intronic (c.832 + 36A>G, c.833 - 29C>T, c.1066 + 82C>G) and 3' UTR (c.*10C>T, c.*41G>A) in a cohort of 5 Italian patients with FALS [302]. Vance conducted a survey of 197 familial ALS British Kindred cases and identified the same R521C mutation in four additional families, as well as two additional missense mutations (R521H, R514G) in another four families [303].

In 2010, Baumer sequenced all 15 exons of the FUS gene in 3 patients revealing a novel deletion mutation (c.1554_1557delACAG) in 1 individual and the c.1574C>T (P525L) mutation in 2 others [304]. Belzil amplified all 15 exons of FUS, sequenced in 154 unrelated FALS cases and 475 ethnically matched healthy individuals identifying a new splicing mutation (R514R) in the highly conserved C-terminal of the FUS protein [305]. Blair examined 247 SALS and 124 FALS Australian patients with European ethnicity and discovered R521C and R521H mutations [306]. Bosco found a novel R495X mutation in a family when 2 patients' DNA was examined [307]. Broustal found R521H mutation in a French ALS patient [308]. Groen studied 52 unrelated Netherland FALS patients observing R521C, R521H, a novel S462F and a polymorphic Q210H mutations [309]. Hewitt screened an initial cohort of 42 familial ALS (FALS) and 117 sporadic ALS (SALS) cases with subsequent larger cohort of 431 SALS cases. Four heterozygous mutations (Gly507Asp, Arg524Trp, Gly174del, Gly228_Gly229insGly), 1 of which is novel, were identified by Corrado in 6 patients with ALS (4 with FALS and 2 with SALS) [299]. DeJesus-Hernandez identified two novel mutations ((g.10747A>G; IVS13-2A>G) and p.G466VfsX14) in FUS in two out of 99 (2.0%) sporadic ALS patients [310]. Huang found P525L mutation in exon 15 of a case [311]. Millecamps screening a cohort of 162 families enrolled in France and 500 controls identifying 31 pathogenic missense mutations found in 36 patients (20 SOD1, 1 ANG, 1 VAPB, 7 TARDBP and 7 FUS - R521C (3 patients) and R514S, R521S, R521H, R521L) [236]. Rademakers screened FUS in a cohort of 200 ALS patients [32 FALS and 168 sporadic ALS (SALS)] identifying a mutation (R521C) that was also present in her affected daughter and a novel mutation (G187S) in one SALS patient [312]. Suzuki found the R521C FUS mutation in a Japanese family with FALS after screening 40 FALS families in Japan and found 4 more mutations (S513P, K510E, R514S, H517P) in exon 14 and 15 of FUS [313]. Tsai screened a cohort of 15 index patients of Taiwanese Han Chinese descent with adult-onset FALS identifying 7 different mutations in 8 patients, including 3 in SOD1, 2 in exon 15 of FUS (H517D and R521H), and 2 in exon 6 of TARDBP [242]. Van Langenhove performed mutational analysis of FUS in 122 patients with FTLD and 15 patients with FTLD-ALS, 47 patients with ALS and in 638 control individuals in a Belgian population and observed the known R521H mutation in 1 patient with ALS apart from 1 patient with FTLD mutation [314].

Yamamoto-Watanabe examined a Japanese ALS6 family with mutation R521C in the FUS/TLS gene [315]. Waibel analyzed a German amyotrophic lateral sclerosis (ALS) cohort, including 133 patients with sporadic ALS and 58 patients with FALS identifying 2 novel heterozygous FUS/TLS mutations in 4 German ALS families including the novel missense mutation R510K and the truncating mutation R495X [315]. Yan identified a total of 17 FUS mutations(S96del, G174-G175del, G206S, G223-226del, Y485AfsX514, R495X, R495EfsX527, K510WfsX517, R521C, R521G, R521H, R521L, R524S, R525L, 5'-2A>T) including 11 novel mutations in 22 FALS families [316].

In 2011, Belzil tested 475 SALS cases of European origin and 475 matched controls for coding variations in the 15 exons of the FUS gene identifying novel FUS mutations (P18S, N159Y, R383C, R502fsX15, Q519X) in a total of five SALS patients [317]. Chio presented the first case of an ALS patient carrying a de novo missense mutation of the FUS gene (c.1561C>T, p.R521C) [318]. Drepper identified R521C, in one familial and one sporadic ALS patient in Germany [319]. Lai in mutational screening of all coding exons of FUS in a total of 1523 ALS cases having 228 sporadic ALS cases, six variants (R521C, R521H, Y66Y, G507D, R518G, P525P) in six different cases were found [320]. Robertson performed a mutation analysis of FUS in a Canadian ALS patient of Chinese origin which revealed an unusual novel heterozygous double point mutation (R514S/E516V) confirming that exon 15 is a mutation hot-spot [321]. In a Spanish cohort, Syriani screened 25 FALS patients identifying 2 mutations R521C and K510E in 2 patients [322]. H517D and R521H mutations in two Taiwanese patients were found in exon 15 of FUS by Tsai [242].

In 2012, Belzil found a novel 1-base pair deletion frameshift R495QfsX527 in exon 14 in a 19-year old male patient [323]. Brown reported 4 new mutations (Ser439Ser,c.105dup, c.132C>A, c.190C>A) and replicated mutations (Gly144_Tyr149del, Gly174del, Ser439Ser, Arg521Gly, Pro525Leu) in a US ALS cohort [252]. Conte found a juvenile form of ALS mutation P525L in a young Italian girl of 11 years of age [324]. Mochizuki found a replicated mutation P252L in 3 members of a Japanese family [325]. Nagayama discovered a novel mutation p.Met464Ile in a SALS patient with corticobassal degeneration. Sproviero screened 327 Italian patients and found replicated mutations R521G, R521C and P525L in 4 patients [326]. Zou sequenced all exons in 10 FALS, 210 SALS and 151 healthy controls discovering R521L, R521H, P151P, R216R, G228G, Y97Y, G49G, G488G and 14C>T mutations [327]. Van Blitterswijk examined 111 FALS from 97 families, 1192 SALS and 970 controls reporting S115N, Q210H, R487C, R495X, R521H and R521C mutations [328].

1.18.1.6 ALS8 - VAPB

VAMP-associated protein B gene is associated with autosomal dominant form having 6 exons on chromosome 20q13.33 consisting of 347 amino acids with 30 mutations reported so far.

In 2004, Nishimura initially described a single P56S mutation in 12 affected individuals in a Brazilian family of Portuguese descent [329]. Landers screened 80 familial ALS cases and identified an exon 5 three-base pair deletion of the serine residue at position 160 (ΔS160) [330].

In 2005, Nishimura observed the same mutation in seven additional Brazilian families [331].

In 2006, according to Marques, all the affected members tested presented the Pro56Ser mutation on exon 2 of the VAPB gene, which was absent in the remaining 34 non affected members [332]. Conforti suggested that sporadic ALS is not associated with VAPB gene mutations in Southern Italy [333].

In 2010, Funke identified P56S in a German pedigree [334]. Millecamps screening a cohort of 162 families enrolled in France and 500 controls identifying 31 pathogenic missense mutations found in 36 patients (20 SOD1, 1 ANG, 1 VAPB - P56S, 7 TARDBP and 7 FUS) [236].

1.18.1.7 ALS9- ANG

Angiogenin gene is a causative gene for a juvenile autosomal recessive form of motor neuron diseases (MNDs) having 33 introns and 2 exons on chromosome 14q11.1 consisting of 1657 amino acids with 18 mutations reported so far.

In 1999, Hayward identified three rare polymorphisms in the untranslated region of the APEX nuclease gene and one common two-allele polymorphism (D148E) on chromosome 14 [335].

In 2004, Greenway identified a novel mutation in two individuals with sporadic ALS that potentially inhibits angiogenin function thereby propose a candidate region for ALS on chromosome 14q11.2 [336]. Skvortsova also analyzed M235T polymorphism of ANG gene as part of his study [337] .

In 2006, Greenway reported the finding of seven missense mutations (Q12L, K17E, K17I, R31K, C39W, K40I and I46V) in 15 individuals, of whom four had familial ALS and 11 apparently 'sporadic' ALS after genotyping the rs11701 SNP in 1,629 individuals with ALS and 1,264 controls from five independent populations. This confirmed the association in the Irish and Scottish populations with ALS, although no association was observed in the populations from the US, England or Sweden [338].

In 2007, Crabtree further showed that the ribonucleolytic activity of 6 of the 7 ANG-ALS implicated variants is significantly reduced or lost and some variants also show altered thermal stability [339]. Corrado screened 262 Italian SOD1 negative ALS patients (250 sporadic) and 415 matched controls for sequence variations in the regions of the ANG gene and did not detect the association with rs11701 in the Italian population [340]. Subramanian identified Q12L, C39W and K40I [341]. Wu sequenced the coding region of ANG in an independent cohort of North American ALS patients and identified four mutations which are associated with functional loss of ANG activity (P(-4)S, S28N, and P112L are novel, and K17I previously reported) in the coding region of ANG from 298 ALS patients [342]. In 2008, Conforti investigated the role of ANG gene in ALS patients from southern Italy and found a novel mutation (in the signal peptide of the ANG gene in a sporadic patient with ALS supporting the evidence that the ANG gene is involved in ALS [343]. Del Bo also argued against the hypothesis of both genes as risk factors for motoneuron neurodegeneration, at least in an Italian population [344]. Gellera collected peripheral blood samples from 737 Italian patients (455 males and 282 females) identifying seven different mutations, five of which are novel (g.446C→T, F-13S, G20G, V113I and H114R), in nine patients (six SALS and three FALS), but not in 515 healthy controls. These findings however suggested the hypothesis that ANG gene mutations may not only increase the risk of ALS but may also be pathogenic due to the absence of these mutations in healthy individuals, their overrepresentation in FALS compared to SALS cases in the cohort, and the strong evidence from functional assays that ALS-linked ANG mutations affect the biological properties of the protein raise [345]. Paubel observed a previously identified a heterozygous mutation (pI46V) in 2 patients with ALS without a known family link and found a novel heterozygous mutation (pR121H) in 1 male patient who developed ALS with rapid progression but no association with rs11701 SNP when 855 French patients with sporadic ALS were analyzed. No mutation was observed in 234 healthy control individuals [346].

In 2009, Fernandez-Santiago investigated the role of ANG in the German population, screening for mutations by sequencing the entire coding region of the ANG gene in a large sample of 581 German ALS cases and 616 sex- and age-matched healthy controls and identifying two heterozygous missense variants, F(-13)L and K54E, in two sporadic ALS cases but not in controls. It was suggested that this finding provided further evidence to support the hypothesis that angiogenic factors up-regulated by hypoxia are involved in the pathophysiology of ALS [347].

In 2010, Millecamps screening a cohort of 162 families enrolled in France and 500 controls identifying 31 pathogenic missense mutations found in 36 patients (20 SOD1, 1 ANG - K17I, 1 VAPB, 7 TARDBP and 7 FUS) [236].

In 2011, Luigetti found a novel heterozygous ANG missense variant (c.433 C>T, p.R145C) which was not found in controls [348].

In 2012, Brown identified 2 novel mutations Tyr38His and Asp46Gly [252]. Kirby identified a replicated mutation K54E in a single SALS case when 517 patients and 278 normal controls were screened from North England [349]. Zou sequenced 10 FALS, 202 SALS and 151 healthy controls in a Chinese cohort, discovering a novel variant V103I in a sporadic patient and a novel H84R variant in a healthy control [327].

1.18.1.8 ALS10 - TDP43

TAR DNA-binding protein gene is associated with autosomal dominant form having 6 exons on chromosome 1p36.22consisting of 414 amino acids with 50 mutations reported so far.

In 2007, Dickson suggested that the study of TDP-43 is useful in differentiating FTLD-MND and ALS from other conditions connected with upper or lower motor neuron manifestations [350].

In 2008, Daoud screened the TARDBP gene in 285 French sporadic ALS patients identifying the insertion of an adenine in exon 6 (c.1255-1256insA) that creates a premature stop codon (Y374X) and three heterozygous missense mutations (P363A and A382P and G348C) in five unrelated patients [351]. Gitcho identified A315T mutations segregated with all affected members of an autosomal dominant motor neuron disease family which was not found in 1,505 healthy control subjects [352]. Kabashi reported eight missense mutations in nine individuals—six (D169G, G287S, G348C, R361S, N390D, N390S) from individuals with sporadic ALS (SALS) and three (A315T, A382T -in 2 individuals) from those with familial ALS (FALS) after screening 200 individuals with ALS (80 FALS, 120 SALS) from France and Quebec; and 185 controls matched for age and ethnicity [353]. Kuhnlein screened 134 patients with sporadic ALS, 31 patients with familial ALS, and 400 healthy control subjects identifying 2 missense mutations (G348C and the novel N352S)in 2 small German kindreds [354]. Rutherford analyzed a cohort of 296 patients with variable neurodegenerative diseases associated with TDP-43 histopathology and 3 different heterozygous missense mutations in exon 6 of TARDBP (M337V, N345K, and I383V) were identified in the analysis of 92 familial ALS patients with no mutations detected in 24 patients with sporadic ALS or 180 patients with other TDP-43–positive neurodegenerative diseases neither in 825 controls and 652 additional sporadic ALS patients [355]. Sreedharan screened 154 index FALS cases for mutations in TARDBP identifying a missense mutation (M337V) in exon 6 from a Caucasian family of English

descent [356]. Van Deerlin identified 2 TARDBP variants Gly290Ala and Gly298Ser in two kindreds with familial ALS when 259 patients with ALS, FTLD, or both were sequenced [357]. Winton identified a single living patient with a A90V substitution in exon 3 in FTLD/ALS case [358].

In 2009, Baumer identified in one patient the novel A321G variant, also A315A and A66A which were previously reported in 2 other patients in Oxford [359]. Corrado screened a large cohort of 666 Italian ALS patients (125 familial and 541 sporadic cases) identifying 12 different heterozygous missense mutations (N267S, G294V, G295S, G295R, S332N, G335D, M337V, S379P, S379C, A382T, G287S, S393L) absent in 771 matched controls [360]. Del Bo carried out a genetic analysis data on TARDBP in 314 ALS mainly Italian patients, identifying four heterozygous missense mutations (G348C, A382T, G294V and G295S) in five unrelated ALS patients [361]. Kamada performed mutational screening of TARDBP in 30 SOD1-negative FALS patients and found a N352S mutation in one case of FALS, but no TARDBP mutations were found in cases of SALS [362]. Kirby screened 42 FALS, 9 ALS-FTD, 474 SALS and 45 PMA identifying 4 mutations in a Northern England cohort of which 2 are novel (Ala321Val and Gly348Val) [363]. Lemmens suggested a novel pathogenic missense mutation in exon 6 of TARDBP (M311V) in one patient with autosomal dominant FALS but absent in 601 healthy controls [364]. Ticozzi screened DNA from 397 ALS patients (208 FALS and 188 SALS) cohort revealing 10 different heterozygous missense mutations (of which three are novel - N352T, N378D and G384R, and others previously reported are G295R, A315T, N345K, G348C, S379P, I383V, N390D) in 14 patients [365].Williams conducted a mutation analysis of TARDBP in an Australian cohort of 74 sporadic and 30 familial ALS cases identifying a novel familial ALS mutation G294V [366].

In 2010, Iida screened the TARDBP mutation in 721 Japanese ALS by direct sequencing identifying a novel mutation (Gly357Ser) and a known mutation (Asn352Ser) in sporadic ALS [367]. Millecamps screened a cohort of 162 families enrolled in France and 500 controls identifying 31 pathogenic missense mutations found in 36 patients (20 SOD1, 1 ANG, 1 VAPB, 7 TARDBP – G295S, A315T, G348C, A382T (2 patients), G384R, W385G and 7 FUS) [236]. Origone discovered S393L variant in an Italian cohort [237] and was replicated in a study by Praline in an ALS family[368]. Tsai screened a cohort of 15 index patients of Taiwanese Han Chinese descent with adult-onset FALS identifying 7 different mutations in 8 patients, including 3 in SOD1, 2 in exon 15 of FUS and 2 in exon 6 of TARDBP (M337V and N378D) [242]. Xiong performed a direct sequencing of 71 SALS patients and 5 FALS Chinese families identifying a novel heterozygous variation, Ser292Asn but not found in 200 unrelated control subjects. Also, two silent mutations (Gly40Gly and Ala366Ala) and one novel polymorphism (239-18t>c) were found [369]. Yokoseki found a Q343R substitution in three affected individuals in two generations of one family [370].

In 2011, Borghero discovered a 53 year-old patient carrying a homozygous A382T mutation [371] supporting the publication by Chio who found a heterozygous A382T mutation in 39 patients when 135 Sardinian ALS patients and 156 controls were screened [372]. Fujita identified A315E mutation in one FALS patient [373] while Ticozzi screened exon 6 in 3 Italian cohorts (376 AD, 463 PD and 376 ALS patients) discovering two novel mutations N352T and G384R [374]. De Marco conducted a study identifying G368S and A382T mutations [375].

In 2012, Brown identified two previously reported mutations Gly348Val and Ala382Thr amongst other mutations in other genes [252]. Chiang, Lida, Lattante, Mosca replicated the homozygous A382T mutation in 2 subjects in a family of 2 generations [376]. Piaceri found a single SOD1 mutation and 3 TARDBP mutations (Ala382Thr, Gly295Ser, Gly294Val) in 5 SALS patients from 61 Italian ALS patients [377]. Solski found a novel insertion/deletion (indel), c.1158_1159delAT; c.1158_1159insCACCAACC heterozygous mutation [256]. Zou examined 312 SALS, 13 FALS and 245 healthy controls in a Chinese cohort identifying 2 SALS patients with the heterozygous S292N, 1 SALS patient with heterozygous G348V and 2 SALS patients with A366A [378].

1.18.1.9 ALS11 - FIG4

FIG4 homolog, SAC1 lipid phosphatase domain containing (S. cerevisiae) gene is associated with autosomal dominant form having 23 exons on chromosome 6q21 consisting of 347 amino acids with 10 mutations reported so far.

In 2009, Chow identified ten unique non synonymous variants of FIG4 in nine patients, including six with SALS and three with FALS (R183X, Q403X, D53Y, D48G, R388G, I411V, Y647C, I902T, R23del(23-55), S424del insR) [379].

1.18.1.10 ALS12 - OPTN

Optineurin gene is a causative gene of hereditary primary open-angle glaucoma for a both recessive and dominant traits having 33 introns and 34 exons on chromosome 10p13 consisting of 1657 amino acids with 23 mutations reported so far.

In 2010, Maruyama investigated six Japanese individuals and found three types of OPTN mutations in three patients: a homozygous deletion of exon 5, a homozygous Q398X nonsense mutation and a heterozygous E478G missense mutation within its ubiquitin-binding domain confirming that these findings strongly suggest that OPTN is involved in the pathogenesis of ALS [380]. Belzil evaluated OPTN mutations in a sample of familial and sporadic ALS cohorts in 95 unrelated familial ALS (FALS), 95 sporadic ALS (SALS) cases and 190 control participants of European descent discovering two newly identified variants (p.K59N and c. 1242 _

1G _ A_insA) and a previously unreported missense p.A481V in 3 individual FALS cases but not identified in any SALS patients [381].

In 2011, Del Bo screened a group of 274 ALS Italian patients, including 161 familial (FALS) and 113 sporadic (SALS) cases which revealed six novel heterozygous variants in both FALS and SALS patients (three missense - p.T282P, p.Q314L, p.K557T, one nonsense p.G23X and two intronic mutations c.552+1delG, c.1401 +4A/G) but absent in 280 Italian controls and over 6800 worldwide glaucoma patients. Among the six novel variants identified in our cohort, the missense p.Q314L and p.K557T mutations were predicted to be pathological and probably affecting protein function by three distinct prediction softwares (PMUT, PolyPhen and SNPs3D), while the p.T282P was scored as neutral by the same programs thereby supporting the possible pathological role of optineurin in motor neuron disease, but not in FTD [382]. Iida screened the entire coding region (exons 4-16) and exon-intron boundaries of OPTN mutation in 713 ALS (687 sALS and 26 fALS) patients to validate previous result using a different cohort and found 17 kinds of sequence variations of which two are novel (p.A93P and p.R271C) and not observed in 940 controls. PolyPhen and the SMART programs predicted that p.Ala93Pro has a possibly damaging function against the gene product while p.R271C was benign and not evolutonarily conserved thereby concluding that OPTN mutation is not a common cause of ALS in Japanese population [383]. Millecamps sequenced the coding exons of OPTN in 126 French patients with familial ALS (FALS) identifying heterozygote nonsense 691_692insAG and p.Arg96Leu variant in a family with dominant ALS [384]. Sugihara screened 563 sporadic ALS (SALS) subjects and 124 familial ALS (FALS) subjects who were mainly Caucasian and found a new c.964T_C synonymous variation (p. S218S) in exon 8 apart from some previously reported variations. It was concluded that further pathological or biochemical investigation studies are needed to show that OPTN may be involved in the pathogenesis of a wide range of ALS [384]. van Blitterswijk identified 1 novel nonsense mutation (Q165X) and 1 unreported missense mutation (Q454E) in individual SALS patients after screening a large Dutch cohort of 1191 patients with SALS, 94 patients with FALS, and 1415 control subjects for mutations and SNPs in OPTN. G159V and Q454E are predicted as probably damaging and possibly damaging respectively [385]. Tumer investigated 64 Danish SALS and FALS cases for mutations in the OPTN gene and identified a heterozygous nonsense p.Gln165X mutation in exon 6 segregating in a family with ALS but not present in 1070 healthy Danish controls and 1000 Danish individuals with metabolic phenotypes or among 3928 glaucoma patients (2803 non-Japanese and 1125 Japanese) and 2943 controls (2340 non-Japanese and 603 Japanese) [386] as reviewed by Maruyama [380].

1.18.1.11 ALS13 – ATXN2

Ataxin 2 gene is significantly associated with ALS having intermediate-length polyQ expansions (27–33 glutamines). It is situated on chromosome 12q23-q24.1 with a record of 1 mutation so far.

In 2010, Elden discovered that ATXN2 were significantly associated with ALS from studying 915 patients [387].

In 2011, 1294 European ALS patients and 679 matched healthy controls were examined by Lee in an European cohort [388]. Also, in another study by Van Damme, 1948 SALS, FALS cases and 2002 controls from Belgium and Netherlands were analysed [389].

1.18.1.12 ALS14 - VCP

Valosin-containing protein is a gene on chromosome 9p13 with 7 mutations reported so far.

In 2010, Johnson sequenced 210 cases from unrelated families and 78 pathologically proven ALS cases, and identified four mutations in the VCP gene (R191Q, R159G, R159C, D592N) found in three affected individuals within an Italian family, and a fourth affected member was an obligate carrier. None of these mutations were found in 569 US control samples, in 636 Italian control samples, in 364 African and Asian samples.

In 2012, Kopper screened a cohort of 93 FALS, 754 SALS, 58 SALS-FTD, and 264 PMA patients discovering a replicated mutation R159H in a FALS patient and .I114V in 1 SALS patient [390]. Miller analysed British FALS and SALS patients but failed to identify an exonic variant in the subset [391].

1.18.1.13 ALS15/UBQLN2

Ubiquilin 2 is a gene on chromosome Xp11.21 with 6 mutations reported so far.

In 2011, mutations in UBQLN2, which encodes the ubiquitin-like protein ubiquilin 2, cause dominantly inherited, chromosome-X-linked ALS and ALS/dementia [63].

1.18.1.14 ALS16/SIGMAR1

Sigma non-opioid intracellular receptor 1 is a gene on chromosome 9p13 with 1 mutation reported so far.

In 2011, the role of sigma-1 receptors in motor neuron function and disease was described in a consanguineous family segregating juvenile ALS in an autosomal recessive pattern and how the genetic variant is responsible for the disorder [64].

1.18.1.15 ALS17/ CHMP2B

Chromatin modifying protein 2B is a gene on chromosome 3p12.1 with 2 mutations reported so far.

In 2006, ALS patients were screened for the CHMP2B gene in chromosome 3p12.1 which is associated with frontotemporal dementia and two mutations were discovered in two patients [65].

1.18.1.16 ALS18/PFN1

profilin 1 is a gene on chromosome 17p13.3 with 4 mutations reported so far.

In 2012, two large ALS families were exome sequenced with results suggesting that mutations within the profilin 1 (PFN1) gene can cause FALS [66].

1.18.1.17 ALS19/ ERBB4

v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 4 is a gene on chromosome region 2q33.3-q34 with 2 mutations reported so far.

In 2013, a whole genome sequencing of an autosomal-dominant mode of inheritance with incomplete penetrance revealed the mutations c.2780G>A (p. Arg927Gln) and c.3823C>T (p. Arg1275Trp) in ERBB4 [67].

1.18.1.18 ALS20/ HNRNPA1

Heterogeneous nuclear ribonucleoprotein A1 is a gene on chromosome 12q13.1 with 1 reported so far.

In 2013, heterogeneous nuclear ribonucleoproteins was discovered in one case of familial ALS [68].

1.18.1.19 ALSFTD1 – C9ORF72

C9orf72 gene is located on chromosome 9p21.2 with hexanucleotide repeat expansion causing amyotrophic lateral sclerosis (ALS) and/or frontotemporal dementia (FTD).

Efforts made towards discovering a link between ALS and Chromosome 9 in the year 2000 was when Hosler conducted a genome wide linkage analysis discovering that a defective gene on chromosome 9 may be linked to amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD) [273] while In 2003, Ostojic studied a Swedish ALS-FTD pedigree in 17 individuals in 2 generations but found no evidence of linkage to chromosome 9q21-22 [392].

In 2006, Morita confirmed a link on chromosome 9p21.3-p13.3 to ALS-FTD when a genome wide linkage analysis was performed on 50 members of a Scandinavian family [393] . Vance also conducted a genome wide linkage study on a large ALS-FTD kindred using Affymetrix microarrays[394]

In 2007, Valdmanis identified 3 ALS-FTD families from Canada and France showing evidence on linkage to chromosome 9p [395].

In 2008, Luty discovered a genetic linkage on chromosome 9p21 from FTD, MND and TARDBP patients [396].

In 2009, Le Ber tested 10 French families linking them to chromosome 9p [397] while van Es screened 2323 SALS and 9013 controls on regions chromosome 19p13.3 and 9p21.2 detecting a high linkage to SNP rs12608932 [398].

In 2010, Boxer conducted a study on chromosome 9p-linked FTD-ALS family to find features of affected patients [399]. Gijselinck conducted a genome wide linkage study of loci at chromosomes 9 and 14 on a Belgian family and found a significant evidence of a linkage to chromosome 9p23-q21 [400]. Laaksovirta conducted a genome wide association study on a Finnish population sampling 442 ALS patients and 521 control individuals and concluding that chromosome 9p21 is a major cause of FALS [401]. Pearson studied a large family from Gwent in South Wales identifying a large haplotype on chromosome 9p21 [402]. Shatunov found a strong evidence of two SNPs (rs3849942 and rs2814707) associated with SALS by conducting a genome wide association study on SALS and controls from 20 UK hospitals which included 599 patients and 4144 controls [403].

In 2011, DeJesus-Hernandez and Renton in their independent studies found expanded GGGGCC Hexanucleotide Repeat in the C9ORF72 gene as the major cause of FTD and ALS while Renton specifically discovered that one-third of European descent FALS have HRE in C9orf72 making it the most common cause of ALS [404, 405]. Iida conducted a replication study on 2 SNPs: rs2814707 on 9p21.2 and rs12608932 on 19p13.3 previously found in SALS Caucasian patients and concluding that no association exists in SALS East Asian patients (comprising of Japanese and Chinese) [406].

In 2012, Andersen, Bigio, Murray, Wood and Orr wrote to confirm the association of HRE of C9orf72 with ALS patients [407-412]. Boeve screened a total of 53 subjects from 43 families to determine the characteristics of patients with repeat expansion in C9orf72 [413]. Brettschneider performed genetic analysis on autopsy confirming the association of hexanucleotide repeat expansions with ALS-FTD patients [414]. Byrne examined the characteristics of ALS patients from the population register collated since 1995 in Ireland [415] while Coon also researched the features in a Scandinavian kindred [416]. Calvo also confirmed the association of GGGGCC expansion in ALS-FTD patients [417]. Chester suggested that bulbar progression in C9orf72 HRE is common and destructive in Portuguese patients [418] while Chio discovered 2 Sardinian families with C9orf72 and TARDBP mutations causing a rapid degeneration [419]. Cooper-Knock studied the features of

ALS-FTD patients with expansions in C9orf72 in a Northern England cohort of 563 cases [420]. García-Redondo assessed 781 SALS, 155 FALS, 248 controls from a Spanish cohort and from various populations globally as reported in HapMap [421]. Gijselinck studied 465 patients with ALS and/or FTD and 856 controls from Flanders-Belgium identifying a HRE in C9orf72 [400]. Herdewyn conducted a whole genome sequencing on 5 members of a family with classical ALS confirming repeat expansion as a cause of FALS [422]. Irwin also examined ALS-FTD patients with HRE in C9orf72 confirming a severe rate of cognitive decline in affected patients [423]. Ishiura analysed a total of 310 ALS and PDC patients from Kii peninsula of Japan and its environs concluding that a severe prevalence of the disease leads to HRE in ALS patients of Kii population [424]. Jang investigated 8 FALS and 246 SALS Korean patients but found no patient with expansion thereby concluding that repeat expansion is not a common cause of ALS in Korean patients [425]. Konno found 2 FALS patients in 58 cases examined which suggests that HREs exist in a Japanese cohort but in small proportion [173]. Koppers examined 1019 SALS and 1103 controls from Dutch cohort by analysing mutations in chromosomes 9p21.2 and 19p13.3 concluding that variants in these coding regions are rare in ALS Dutch patients [390]. Lattante identified C9orf72 patients apart from 6 other genes screened in an Italian cohort of 480 SALS and 48 FALS patients [426]. Majounie screened 1425 FTD patients and 4448 ALS patients confirming that a common C9orf72 gene is involved in many cases of familial and sporadic ALS and FTD [427]. Millecamps studied 950 French ALS patients and 580 controls confirming that expanded repeats in C9orf72 plays a vital role as a causative for ALS [428]. Mok discovered that expansions in the C9orf72 gene is a common cause of ALS in the Greek population [429]. Ogaki found no evidence of C9orf72 in FTLD Japanese patients [430]. Pamplett studied 43 Australian SALS patients and found expanded repeats in patients [431]. Rademakers found hexanucleotide repeat expansion in ALS-FTD patients [432]. Ratti found a hexanucleotide repeat expansion when a large cohort of 259 familial ALS, 1275 sporadic ALS, and 862 control individuals of Italian descent were screened [433]. Rutherford identified HREs in patients with ALS [434]. Sabatelli reported HREs as the most common mutation in a Sardinian and Italian SALS population [435]. Savica examined the behavioural pattern of 3 ALS-FTD subjects [436]. Simón-Sánchez investigated a Dutch cohort of 353 FTDsporadic or familial patients with or without ALS [437]. Smith Stewart found HRE mutation in 17FALS and 6 SALS when 231 ALS patients were examined [438].

In 2013, Beck discovered HREs in multiple neurodegenerative diseases in the UK population [439]. Dombroski studied Guam patients with Amyotrophic Lateral Sclerosis-Parkinsonism-Dementia Complex in the Chamorros population suggesting that C9orf72 expansion is not a common cause of ALS-PDC [440]. Gomez-Tortosa confirmed an association of C9orf72 HREs with ALS-FTD patients studied in a Spanish cohort [441]. Jones

performed a genome wide association testing on ALS samples and found SNPs rs3849942 and rs903603 strongly associated with ALS [442]. Zou screened the HRE in C9orf72 in a Chinese ALS cohort of 324 SALS and 20 FALS but could not identify an expanded hexanucleotide repeat [443].

1.18.1.20 ALSFTD2 – Unknown

In 2013, Dobson-Stone reported a locus on chromosome 16p12.1-q12.2 in a genetic analysis on a large European Australian kindred [440].

1.18.1.21 OTHERS

1.18.1.21.1 DAO

D-amino acid oxidase gene is associated with classical adult onset familial amyotrophic lateral sclerosis (FALS) having 10 exons on chromosome 12q24 consisting of 347 amino acids with 2 mutations reported so far.

In 2010, Mitchell reported an exon 7 heterozygous variant R199W in a UK three generational FALS kindred while Millecamps identified an exon 2 heterozygous variant R38H in a French woman [444, 445].

1.18.1.21.2 DCTN1

Dynactin gene is associated with autosomal dominant form and 31 exons on chromosome 2p13 consisting of 347 amino acids with 2 mutations reported so far.

In 2003, Puls identified G59S variant in a North American family with a slowly progressive, autosomal dominant form of lower motor neuron disease without sensory symptoms [446].

In 2004, Munch screened 250 patients with ALS and 150 unrelated controls identifying three mutations T1249I, M571T and R785W [447].

In 2005, Munch reported another heterozygous R1101K mutation in a family with ALS [448].

In 2008, Takahashi identified a novel putative pathogenic R997W mutation in exon 18 apart from other novel variants in SOD1, ALS2, ANG and VEGF [225].

1.18.1.21.3 NEFH

Fusion gene is associated with autosomal dominant form having 3 introns on chromosome 22q12.2 consisting of 347 amino acids with 30 mutations reported so far.

In 1994, Figlewicz found some deletions within the C-terminal KSP repeat region of the neurofilament heavy-subunit gene some human ALS patients [449]

In1996, Rooke found no variation in the NEFH gene in 117 unrelated cases of familial ALS when the C-terminal KSP repeat region was examined [450] and Vechio also found no deletions in the NEFH gene [451].

In 1998, Tomkins identified an 84-bp insertion in the NEFH tail in a study of 164 ALS cases and 207 age-matched controls [452].

In 1999, Al-Chalabi analyzed samples from 2 different populations (UK, 207; Scandinavia, 323) with age-matched controls for each group (UK, 219; Scandinavia, 228) and found 4 novel NEFH tail deletions [453].

In 2004, Skvortsova examined ALS patients from a Russian cohort suggesting that the S genotype of the gene is associated with sporadic MND in patients in Moscow [337]

In 2011, Daoud screened 190 ALS patients sequencing 29 candidate genes and discovered 40 novel variants including NEFH variants –Lys867Asn, Glu918Gly, Arg346His, Ala40Val [454].

1.18.1.21.4 KIFAP3

In 2009, Landers conducted a genome wide association study on 1821 SALS and 2258 controls from US and Europe yielding the SNP rs1541160 with a pvalue of 1.84 x 10(-8) [455].

In 2010, Traynor made an attempt to replicate the study in 504 Italian ALS patients but the gene had no effect on survival in that population [456].

1.19 Databases

1.19.1 Data

The word data is a plural of ‘datum’ – a latin word meaning ‘to give’. It originated from the pre-computer days of ‘trade by barter’ where goods were measured, exchanged and stored. Data in our modern days is a piece of information which could either be quantitative or qualitative. A collection of the set of variables or data requires organization, storage and availability for use when needed. This therefore invented the idea of the term ‘database’ [457].

1.19.2 Evolution of Database

At the start of early civilization, medical records, government documentations, library materials and businesses required means of storing and retrieving information but only got easier, cheaper and more compact at the advent of computers in the sixties [457]. Navigational database system was the first generation of database in which records were retrieved through navigation from one record to another. The two models were network model and hierarchical model [458].

In 1970, Edgar F. Codd conceived the idea of a relational database [459] (searching for data by content rather than through links) which is different from the previous models. Relational database consists of a schema (structure) and a table (rows and columns) with records linked together with a 'key'.

1.19.3 Relational model: review

A database is a collection of relations (or tables); Each relation has a list of attributes (or columns); Each attribute has a domain (or type) and Each relation contains a set of tuples (or rows) [460, 461].

1.19.4 Database Management Systems

DataBase Management System (DBMS): a software system that facilitates the creation and maintenance and use of an electronic database [458]. The main roles are to keep data around (persistent), answer queries (questions) about data and update data administratively.

1.19.5 Examples of Databases

Oracle, IBM DB2 (from System R, System R*, Starburst), Microsoft SQL Server, NCR Teradata, Sybase, SAP, SQLite, dBASE, FoxPro, IBM DB2, LibreOffice Base, FileMaker Pro, Informix (acquired by IBM), PostgreSQL (from UC Berkeley's Ingres, Postgres), Tandem NonStop (acquired by Compaq, now HP), MySQL, Microsoft Access.

1.20 Genetics database

A genetics database can be simply defined as a structured and comprehensive collection of data on genes, heredity, and variation of organisms [462].

1.20.1 History of Genetics Databases

1949 saw a description of the first mutation that caused human disease, in the haemoglobin gene. Many more mutations led to the need to track these. In the1970s, Prof Victor McKusick rose to this challenge by creating the first database of mutations and inherited diseases, known as Mendelian Inheritance in Man. A transformed

version of this database was later developed for free and easy access online, and is now known as OMIM [463].

The development of an accurate, integrated, complete and systematic database by some geneticists became essential in 1994 to attend to the issue of mutation documentation. A need arose for a regularly updated list of gene mutations accessible online and gene-specific databases integrated with central databases where a systematic mutation naming convention based on amino acid change were agreed [464, 465].

Mutation databases of human genes are becoming more prominent and indispensable in all areas of health care. Also, increasing number of experts in the mutations and diseases of particular genes are curating published and unpublished mutations in locus-specific databases (LSDB) extensively [466]. Apart from being considered as tools that provide vital information on protein function and structure, databases of mutations causing Mendelian disease play a fundamental role in research, diagnostic, and genetic health care [467, 468]. It is therefore pertinent to keep these databases updated and relevant to daily research work.

With the advent of high-throughput sequencing technologies which provide genome sequencing within a shorter time at less cost but with terabytes of data produced. Management of these data could pose storage and administrative problems if not properly carried out [469]. Large-scale Genome Wide Association Studies reveal more genes associated with diseases like ALS. The administrator of the database is therefore posed with the challenge of keeping users of the database informed of new information in the field of ALS.

1.20.2 Non-locus specific mutation databases

Considering the wide range of large databases freely available online, I compiled a list of a few relevant large databases from Pubmed and Google Scholar using the keyword “Databases”. Databases with broken links and locus-specific mutation databases and were excluded from the list as seen in Table 2.

Table 2: Non-locus specific mutation databases

Databases	Full Name	Website	Support
ALFRED	Allele Frequency Database	http://alfred.med.yale.edu/alfred/index.asp	U. S. National Science Foundation
ExPASy	Expert Protein	http://www.expasy.org/	Swiss Institute of

	Analysis System		Bioinformatics
HGMD	Human Gene Mutation Database	http://www.hgmd.cf.ac.uk/ac/index.php	Institute of Medical Genetics Cardiff
LOVD	Leiden Open-source Variation Database	http://www.LOVD.nl/	European Community's Seventh Framework Programme, Netherlands
HGVbaseG2P	The Human Genome Variation Society	http://www.hgvbaseg2p.org now redirected to: http://www.gwascentral.org/	GlaxoSmithKline, the University of Leicester, and the European Community's Sixth Framework Programme and Seventh Framework Programme
HapMap	The International HapMap Project	http://www.hapmap.org	funding agencies from Canada, China, Japan, Nigeria, the United Kingdom and the United States
dbSNP	Human single nucleotide polymorphism (SNP) Database	http://www.ncbi.nlm.nih.gov/SNP/	National Center for Biotechnology Information, U.S. National Library of Medicine

SeattleSNPs	UW-FHCRC Variation Discovery Resource	http://pga.gs.washington.edu/	the National Heart Lung and Blood Institute's (NHLBI) Programs for Genomic Applications (PGA)
PicSNP	Catalogue of non-synonymous SNPs	http://plaza.umin.ac.jp/~hchang/picsnp/	University hospital Medical Information Network Japan
OMIM	Online Mendelian Inheritance in Man	http://www.ncbi.nlm.nih.gov/omim/ or http://omim.org/	Institute of Genetic Medicine, Johns Hopkins Medicine and National Human Genome Research Institute

1.20.2.1 ALFRED

ALFRED (ALlele FREquency Database) is a resource of gene frequency data on human populations supported by the U. S. National Science Foundation. ALFRED has been designed to make allele frequency data on anthropologically defined human population samples readily available to the scientific community and to link these polymorphism data to the molecular genetics-human genome databases. Due to the large volume of data available from the human genome project (HGP) to study human variation, collecting the resulting data into a single, easily accessible resource is vital to facilitate this research [470-472].

1.20.2.2 ExPASy

The ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB) is dedicated to the analysis of protein sequences and structures.

1.20.2.3 HGMD

The HGMD (Human Gene Mutation Database) maintained in Cardiff. HGMD represents an attempt to collate known (published) gene lesions responsible for human inherited disease. HGMD comprises various types of

germ-line mutation within the coding, splicing and regulatory regions of human nuclear genes. As at December 2007, HGMD contained 76,011 different mutations in 2876 human genes [473, 474].

1.20.2.4 LOVD

LOVD is a platform-independent Web-based LSDB-in-a-Box package which lists 71 public LOVD installations hosting 3,294 gene variant databases with 199,000 variants in 84,000 patients.[475]

1.20.2.5 HGVbaseG2P

The Human Genome Variation database of Genotype-to-Phenotype information (HGVbaseG2P) provides a centralized compilation of summary level findings from genetic association studies, both large and small. It is built upon a basal layer of Markers that comprises all known SNPs and other variants from public databases such as dbSNP and the DBGV [476, 477].

1.20.2.6 HapMap

The goal of the International HapMap Project is to develop a haplotype map of the human genome, the HapMap, which will describe the common patterns of human DNA sequence variation. The HapMap is expected to be a key resource for researchers to use to find genes affecting health, disease, and responses to drugs and environmental factors [478, 479].

1.20.2.7 dbSNP

Large-scale sampling designs using association studies, gene mapping and evolutionary biology led to the idea of developing a general catalog of genome variation now known as the dbSNP database by the National Center for Biotechnology Information (NCBI) [480].

1.20.2.8 SeattleSNPs

The SeattleSNPs PGA is focused on identifying, genotyping, and modelling the associations between single nucleotide polymorphisms (SNPs) in candidate genes and pathways that underlie inflammatory responses in humans [481].

1.20.2.9 PicSNP

PicSNP is a catalogue of non-synonymous SNPs (Single Nucleotide Polymorphism) in the human genome. SNPs were retrieved from annotations in the draft sequence available in GenBank. Current version of PicSNP is based on the GenBank database of Feb 6, 2002. Details of assembly of the catalogue is described in reference [482].

1.20.2.10 OMIM

Mendelian Inheritance in Man (MIM) is a brief but comprehensive bibliographic material with examinations on inherited disorders and genes maintained by geneticists and molecular biologists. Its online version, Online Mendelian Inheritance In Man (OMIM), is freely available on the World Wide Web. Unlike other databases that preserve primary sequence, mapping, or reference material, OMIM offers authoritative free text overviews about genetic disorders and gene loci that can be used by researchers [483, 484].

1.20.3 ALS-specific mutation databases

These are disease databases with particular reference to Amyotrophic Lateral Sclerosis only. Apart from ALSoD, the following are independent ALS databases:

Table 3: ALS-specific mutation databases

Databases	Website	Support	Last Update as at 28/02/2014
ALS Mutation Database	https://reseq.lifesciencedb.jp/resequence/SearchDisease.do?targetId=1	Tokyo Univ. Hosp., Tokyo Univ., and Hitachi Ltd.	18/11/2010
Amyotrophic Lateral Sclerosis Gene (ALSGene)	http://www.alsgene.org	Max-Planck Institute for Molecular Genetics, Berlin, Germany with Prize4Life	12/02/2014

1.20.3.1 ALS Mutation Database

Recently, an ALS mutation database constructed as part of the Life Science Integrated Database Project conducted by the Japan Ministry of Education, Culture, Sports, Science, and Technology was published. It contains their original experimental results and published data extracted from scientific journals. The database is expected to play a complementary role to the ALSoD database especially in collecting variations in the Asian region [485].

1.20.3.2 ALSGene

The ALSGene database provides an unbiased regularly updated results of genetic association studies in ALS [486].

1.21 Collection

1.21.1 Data mining, Text mining and Web mining

The advancement in data generation using innovative technologies like high throughput sequencing for GWAS studies and whole exome sequencing has provoked an enormous volume of genetic data by researchers. Methods and techniques development are required for discovering patterns in huge datasets with the aim of extracting sensible information and this process is known as KDD (Knowledge Discovery in Databases). There is a great increase in interest of extracting and mining information from literature due to the free availability of electronic publications in large databases like PubMed [487-489].

Text mining is the process of extracting quality information from text. This is utilized in spam filters to define the pattern of advertisements or spam emails or other unsolicited materials [490-494]. In the field of biological research, the National Centre for Text Mining (NaCTeM - http://www.nactem.ac.uk) – which is a collaboration between University of Manchester and Korea Institute of Science and Technology Information (KISTI) – was publicly funded to conduct research on text mining [495]. BioText is a software developed to assist in analysis and mining of texts [496].

Humans have approximately 24,000 genes [497]. However, communicating and sharing these data is difficult as there is no universal standard terminology to represent biological terms. Gene ontology (GO) therefore helps to provide terms synonymous to a word or string of words covering three domains: biological process, cellular component and molecular function [498-500]. It is what I will refer to as the dictionary of genes.

Most of the bulk of work in text and data mining has already been developed by larger databases like the NCBI, UCSC, HUGO etc., hence the desire not to reinvent the wheel. The motivation for utilizing these larger databases is by manually extracting and storing identity numbers allocated to each gene on ALSoD.

1.21.2 External Databases

Researchers studying ALS genetics are currently obliged to use several major resources (or their equivalents) to find genetic, transcriptomics, protein and clinical information. Examples are PubMed, The HapMap, the 1000 Genomes Project, ArrayExpress, OMIM, Ensembl, the UCSC Genome Browser, Swiss-Prot, EBI and dbGAP. PubMed lists scientific publications and includes links to full articles where available. The HapMap is an international consortium aiming to map all common, bi-allelic human genetic variation in ancestral populations, with the emphasis on Europeans, Chinese/Japanese and Africans. This allows the correlation structure of the human genome to be visualized and for example, makes it simple to develop ancestry informative genetic markers or to use so-called tag SNPs that reduce and simplify genetic association studies.

The 1000 Genomes Project aims to sequence 1000 human genomes to identify rare genetic variation and understand how this relates to human phenotypes. OMIM (Online Mendelian Inheritance in Man) is a searchable database of Mendelian diseases and genes, with referenced reviews that are kept up-to-date. Ensembl and the UCSC Genome Browser are annotated versions of the human and other animal genome assemblies, and include information on gene and genetic variation positions, conservation between species, RNA and protein variation and expression and other information. The exploration of protein findings requires researchers to explore Swiss-Prot or similar protein databases. The European Bioinformatics Institute (EBI) and the genetic association database, dbGAP, are repositories of published genome-wide association data, which researchers have to interrogate to identify negative studies, or to download public genotypes for subsequent analysis.

1.21.3 The need for a Single Resource

Accessing too many data sources or databases could be confusing. What is required is a single resource that can act as a coordinated gateway for the relevant links with a summary, a repository for data that would not be stored elsewhere, and a quick analysis tool that does not require multiple registrations or downloads of huge datasets.

1.22 Integration

Concepts and information are freely shared in the field of science which is the reason for a huge number of curated, up-to-date non-commercial or non-profit databases are freely available [465]. Extracting relevant gene variation information from the massive comprehensive central databases like Genbank, EMBL, Swiss-Prot has become increasingly impossible due to their structure, hence a time-saving and internet-accessible need for locus-specific databases (LSDBs) [465, 468].

ALS genetic data are increasing exponentially with a rising need for combining data existing in different sources to present the ALS-research community with a unified examination of these data [501].

1.22.1 What is needed in ALS

A flawless integration of the two types of online variation databases (which are the genomic and locus-specific) can be achieved with the use of a unique identifier as there exists various ways to create a variation database [465]. Unique ids can be used to systematically link to other bioinformatic tools and comprehensive databases that do not require registration and are freely available online. Apart from NCBI, UCSC, OMIM, KEGG, UNIPROT, GeneMANIA, Genecards, Pubmed, AmiGO, iHop and a host of others, there are other

general information from non-biological databases that could be hyperlinked to have a more general overview of each gene without bias.

1.22.2 Curators/Contributors

Published data and unpublished data curated by the database administrator and registered ALS researchers could be incorporated into the database to avoid unfairness. To keep the database updated with novel variants discovered in various laboratories around the world, I hope to accept and encourage increased participation from more researchers globally.

1.22.3 Google API

Google APIs can display the popularity of a database around the world. It could be used to enable researchers or patients to view the geographical distribution of reported gene variations associated with ALS. Other essential information I hope to make available on the database through data integration from other sources are chromosomal overview of all ALS-related genes, gene overview of all published ALS-related genes, key published studies on all current ALS-related genes from relevant publications and latest news on ALS from various sources automatically made available through Really Simple Syndication (RSS) feeds.

1.22.4 Animal Models

The use of animal models in research helps to provide a distinctive prospect to study incurable and deadly human diseases like ALS both clinically and pathologically by performing studies which are considered unrealistic or impossible to undertake in human patients with motor neuron degeneration [502]. To contribute to a better understanding of the pathogenesis of ALS, links to information on animal models will be a valuable facility on the database. I hope to integrate the dataset into the animal model section of each gene overview as a combination of information from MGI (Mouse Genome Informatics) and Pubmed databases.

1.23 Analysis

The challenge of providing information as required by researchers can only be found in a database that accurately stores, reports and analyses these helpful data [503]. The sequencing of the human genome and the HapMap project [504] have impacted the study of human disease in a significant way and are enabling many genome-wide association studies that aim to elucidate the genetic component of complex diseases [104].

One problem facing researchers, particularly those who may not be geneticists but who rely on genetics data, is knowing how credible any report of linkage or association is. Having an independent measure of credibility

that might guide non-geneticists in knowing the level of confidence to place in reported findings would be an additional useful tool.

1.23.1 Genome Wide Association Study (GWAS)

A major challenge but an extremely desirable outcome would be the online availability of all genome-wide association study (GWAS) data [505] for selective interrogation by users and comparison with other information sources such as linkage, candidate gene or other genome data. GWAS have generated bioinformatics challenges on how to model the complexity of genotype-phenotype connections, integration of biological databases for interpreting genetic associations and the development of a software that enhances collaboration among biologists and bioinformaticians [506]. GWAS is a frequently used method for discovering disease loci by researchers, it is a useful and impartial approach to uncover the risk alleles for genetically complex non-Mendelian disorders. GWAS results are frequently arranged according to level of significance by combining studies via meta-analysis to increase power [507]. Implementing a GWAS interface on a publicly available database for ALS will yield a desirable outcome for researchers to view the results graphically using straight forward queries.

The introduction of GWAS provides the opportunity to cross-examine the entire genome with increased resolution and power. GWAS have already been conducted for a variety of complex diseases, with varying degrees of success. Larger GWAS samples are more likely to discover larger numbers of common susceptibility variants with small effects. A study of controversies and perspectives in the genetics of ALS suggests that human genetic studies in ALS, including GWAS and the study of structural genetic variation, is a strong tool to recognize the molecular pathways involved in ALS. It is necessary to genotype the whole genome both within and between populations as being done by the genome project team in order to fully define the relevance of SNPs to SALS [508].

1.23.2 On-the-fly Meta Analysis

Meta-analysis is the analysis of analyses. A large collection of analysis results from individual studies are statistically analyzed with the intention of amalgamating the findings. A useful tool to perfom a meta-analysis of datasets based on user's defined selection would be desirable. An additional feature to an on-the-fly analytical GWAS tool would be to comapare data from users with available data stored on the database.

1.23.3 Predicting gene interactions

Gene interactions are important for biologists and scientists and would therefore be useful to include. GeneMANIA is an interaction prediction tool described as follows: "the high accuracy of the GeneMANIA

prediction algorithm, an intuitive user interface and large database make GeneMANIA a useful tool for any biologist" [509]. The relevance of a gene interaction tool is indispensable on a freely available genetics database.

1.24 Smartphone Development

Complex diseases like ALS pose a great challenge to researchers worldwide. Research in the field of health sciences is advancing rapidly hence the need for an effective facility that provides storage, analysis, and interpretation of data constantly generated by researchers.

The advent of the internet is one of the useful things that has happened to this generation because it provides efficient and swift online resources within a short time. Apart from social networking, data sharing through websites has widened the horizon of scientific research work. People can now communicate ideas easily to anyone anywhere in the world without any form of physical contact using a computer. Internet services are now available almost everywhere and users access the internet either for fun by playing games or social networking with friends or for more serious work like reading latest news or accessing online resources through websites [510].

Gone are the days of stationery, immovable and gigantic computers that took days to analyze data which were manually punched in using punched tapes. Desktop computers were later introduced as small-sized, movable systems with better input devices like keyboards, mouse and joysticks. Nowadays, Laptop computers are becoming outdated and pocket-sized mobile devices like smartphones and tablets are commonly used to access information online. Most websites were initially built to display valuable information online to users of desktops and laptops. The pages are configured to suit the height and width of the screen perfectly. More people now access websites on various types of devices thereby making it essential to have a usable and effective mobile website that can be displayed correctly on a user's device [511, 512].

Desktop computers were mainly used to browse websites but this has been taken over by mobile devices in the last decade. Data traffic on mobile devices just for the purpose of mobile browsing has risen over four times in 2008 [511]. Apart from making phone calls, mobile phone technology has greatly improved from monochrome screens for sending SMS to colourful graphical touch screens for mobile browsing. [513] The former issues of low bandwidth and low resolution screens on mobile devices are constantly being resolved and due to the regular changes to create a worthwhile browsing experience for mobile device users, smartphones and tablets are getting smarter and are often referred to as "mobile computers" [514-517]. According to Netmarketshare, the introduction of iOS (Apple mobile device operating system), between March

2010 and October 2010 doubled the use of iPad and iPhone for mobile browsing [518] thereby predicting that in 3 years mobile “should” take over desktop internet usage as seen in Figure 1 [519]. This information shows that more users of the ALS database will require access to the database outside the office environment on their mobile devices. A vital tool necessary for a more useful database is to develop a website reachable on any mobile device with an internet connection.

Source: Pingdom, May 2012

This table also shows the dramatic growth rates in mobile share of web traffic across the world over the last two years

Mobile share of web traffic			
	2010	2012	Increase 2010-2012
Africa	5.81%	14.85%	155.59%
Asia	**6.1%**	**17.84%**	**192.46%**
Europe	1.81%	5.13%	183.43%
North America	4.71%	7.96%	69.00%
Oceania	2.88%	7.55%	162.15%
South America	1.46%	2.86%	95.89%
Worldwide	3.81%	10.01%	162.73%

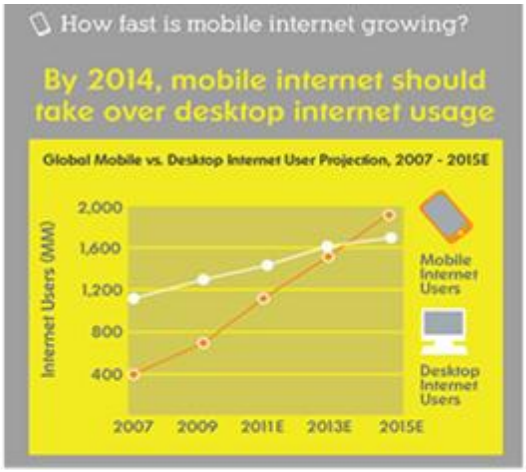


Figure 1: Mobile Marketing Statistics adapted from Netmarketshare webpage

1.24.1 Mobile Websites

Just like any other website, a mobile website has HTML pages which are connected and accessible over the internet (even though mobile websites are typically accessed through WiFi 3G or 4G networks). Mobile websites are designed for smaller handheld and touch screens. Contents of mobile websites like any other website are texts, videos, data and images [520]. Due to the ever-changing technologies for mobile devices, former issues of mobile browsing due to bandwidth has been overcome to a large extent [513].

1.24.2 Apps

These are applications downloaded and installed on a mobile device like a smartphone or tablet instead of accessing the content through a browser. It enables users’ accessibility of contents where there are no connections to the internet. There are various categories of apps downloadable from portals like Blackberry App World, Apple App Store and Android Market [520]. IPhone was first introduced in 2007 thereby leading to the introduction of apps stored in the Apple’s App store for download by iphone and ipad mobile devices.

A comparison between a mobile website and an app is displayed in Table 4with explanation following.

Table 4: Comparison between mobile website and app

Features	Mobile Website	**App**
Instant Availability	Yes	No
Compatibility	Yes	No
Integration	Yes	No
Instant Update	Yes	No
Easily searched	Yes	No
Shareability	Yes	No
Reachability	Yes	No
Permanency	Yes	No
Convertibility	Yes	No
Cost and Time	Yes	No
Support	Yes	No
Speed	No	Yes
Complex reports	No	Yes
Habitual usage	No	Yes
Functionality	No	Yes
Connectivity	No	Yes

Discoverability	No	Yes

1.24.3 Advantages of a mobile website over an app

If the motive behind establishing an extensive mobile presence is just to distribute contents of the website to users and be found on search engines, then a logical option is a mobile website. Here are details of the advantages [520]:

1.24.3.1 Instant Availability

Devices like Blackberry, iPhone and Android allow immediate access to users through a browser whereas, apps need users to first download and install the applications before the content can be viewed.

1.24.3.2 Compatibility

All users are reachable on a single mobile website across various mobile devices whereas apps separate versions are designed specifically for each type of mobile device.

1.24.3.3 Integration

Mobile website URLs are effortlessly incorporated within other mobile technologies like NFC, SMS, QR Codes unlike apps.

1.24.3.4 Instant Update

Changes made to the content of a mobile website are seen immediately whereas an app can only be updated if downloaded by users of each type of device.

1.24.3.5 Easily searched

Mobile websites could be found easily by search engines especially for automatic redirection of regular visitors using handheld devices. But for apps, device manufacturers' app stores restrict the visibility of apps.

1.24.3.6 Shareability

An app cannot be shared easily by publishers within text messages or social network sites or emails using a simple hyperlink like a mobile website.

1.24.3.7 Reachability

An app is only restricted to users of the app-specific device whereas a mobile website can be accessed by a larger number of users across platforms.

1.24.3.8 Permanency

Apps are deleted from users' devices if they are no more relevant to the user thereby having a short lifespan but a mobile website cannot be deleted by a user so they are able to access it when required.

1.24.3.9 Convertibility

A database-controlled mobile website can be extended to act as an app.

1.24.3.10 Cost and Time

Developing a mobile website can be achieved within less time and costs less than having multiple apps across different device platforms.

1.24.3.11 Support

Over time, it is easier to support and sustain a mobile website than an app. Issues of upgrades, compatibility, maintenance and testing are more expensive in apps than in mobile websites.

1.24.4 Advantages of an app over a mobile website

If the aim of developing an app which works more like a computer program instead of a website primarily for interactivity is profit-orientated, then it is more reasonable to develop an app rather than a mobile website. It will be time-wasting, pointless and worthless to embark on an expensive venture of building an app which performs basic functions achievable with a mobile website. Below are more reasons to realize an optimal profit on the investment of app development [520]:

1.24.4.1 Speed

Where Interactivity is essential in accessing data effectively on a mobile device, speed is the main focus.

1.24.4.2 Complex reports

Complex calculations are better performed on apps.

1.24.4.3 Habitual usage

Apps are more efficient at achieving personalized usage of the mobile phone e.g using EverNote on a mobile device is an extremly faster, easier, more proficient way to store and arrange all data, regardless of variety (text, website, video, audio, PDF).

1.24.4.4 Functionality

Apps use the full native functionalities of the handheld device like text messaging, calling, sensors and using the camera more effectively unlike a mobile website.

1.24.4.5 Connectivity

App contents are accessible offline without a need for network or wireless connections unlike mobile websites which can only be made available on a device connected to the internet.

1.24.4.6 Discoverability

Despite charges involved in submitting an app to Android's App Market or Blackberry's App World or Apple's App Store, it is more exciting and offers more publicity when an app is downloadable from the global App market.

1.25 The Project

The title of my project is "**DEVELOPMENT OF TOOLS FOR AUTOMATED COLLECTION, INTEGRATION AND ANALYSIS OF GENETIC DATA IN ALS**".

1.25.1 Problem Statement

The purpose of the research proposed is to develop a dynamic computational web page which tests the hypothesis that a central, automated database can successfully collate the genetic data available and present it in easily accessible forms for researchers. The approach will analyze existing genome-wide association study data that have been collected for meta-analysis and uploaded unpublished user-data. The product of this analysis is an on-the-fly meta-analysis in which unpublished user-data is combined with existing studies confidentially and the result fed back in minutes.

During the course of the research, methods for providing a web-based integration of databases to analyze the function of SNPs found in ALS-related genes will be explored and data will be pooled from various sources like Hapmap, dbSNP etc.

1.25.2 Research Questions

Description of the novel research questions proposed for my thesis project is:

1) Is it possible to generate a database that summarizes genetic data for a disease and allow meta-analysis online? Amyotrophic lateral sclerosis (ALS) will be used as a model disease.

2) Can all genetic data useful to ALS be automatically collected?

3) How achievable is an on-the-fly analysis of genetic data online?

4) To what extent can the database (the ALS Online Genetics Database) be integrated with other bioinformatics resources available online?

5) With the advent of various ALS-related genes, can a database generate levels of evidence to support or refute genetic associations and linkages with ALS?

1.25.3 Overall aim

The overall aim of this proposed research work is to generate and develop a central, automated database that can successfully organise the genetic data available and present it in easily manageable forms for researchers.

1.25.4 Objectives

- To restructure the database by rewriting the schema in order to deal with irregularities and inconsistencies experienced by users.

- The use of up-to-date APIs (Application Programmable Interface) and chart controls compatible with the current webpage application package. This is to present data graphically in a more understandable way.

- To curate data from publicly available databases using automated and manual text mining techniques where necessary.

- To rewrite queries and implement appropriate stored procedures by integrating scripts in SQL (Structured Query Language), javascript, perl language and other relevant scripts required to perform this task.

- To implement more in-house scripts to monitor users visiting the webpage.

- To develop a webpage summarizing, analysing and displaying data from genome wide association studies in ALS.

- To develop a webpage that grabs current news in ALS using RSS (Really Simple Syndication) feeds and other available technology fit for this purpose.

- To explore the use of XML (Extensible Markup Language), Haploview, R, Perl, JavaScript, C#, Visual BASIC integrated under the ASP.NET platform to create a more user-friendly webpage.

Chapter 2 Publication 1 - ALSoD: A User-Friendly Online Bioinformatics Tool for Amyotrophic Lateral Sclerosis Genetics [521]

Contribution: First Author

Journal: Human Mutation

Volume: Volume 33, Issue 9, pages 1345-1351, September 2012

Authors: Olubunmi Abel, John F Powell, Peter M. Andersen, Ammar Al-Chalabi

Pubmed ID: 22753137

DOI: 10.1002/humu.22157

Link:
http://onlinelibrary.wiley.com/doi/10.1002/humu.22157/full

DATABASES

Human Mutation

OFFICIAL JOURNAL HGVS HUMAN GENOME VARIATION SOCIETY www.hgvs.org

ALSoD: A User-Friendly Online Bioinformatics Tool for Amyotrophic Lateral Sclerosis Genetics

Olubunmi Abel,[1] John F. Powell,[2] Peter M. Andersen,[3,4] and Ammar Al-Chalabi[1]*

[1]Department of Clinical Neuroscience, King's College London, Institute of Psychiatry, London, UK; [2]Department of Neuroscience, Institute of Psychiatry, King's College London, London, UK; [3]Section for Neurology, Institute of Pharmacology and Clinical Neuroscience, Umeå University, Umeå, Sweden; [4]Department of Neurology, University of Ulm, Ulm, Germany

For the Databases in Neurogenetics Special Issue

Received 31 January 2011; accepted revised manuscript 19 June 2012.
Published online 2 July 2012 in Wiley Online Library (www.wiley.com/humanmutation).DOI: 10.1002/humu.22157

ABSTRACT: Amyotrophic lateral sclerosis (ALS) is the commonest adult onset motor neuron disease, with a peak age of onset in the seventh decade. With advances in genetic technology, there is an enormous increase in the volume of genetic data produced, and a corresponding need for storage, analysis, and interpretation, particularly as our understanding of the relationships between genotype and phenotype mature. Here, we present a system to enable this in the form of the ALS Online Database (ALSoD at http://alsod.iop.kcl.ac.uk), a freely available database that has been transformed from a single gene storage facility recording mutations in the *SOD1* gene to a multigene ALS bioinformatics repository and analytical instrument combining genotype, phenotype, and geographical information with associated analysis tools. These include a comparison tool to evaluate genes side by side or jointly with user configurable features, a pathogenicity prediction tool using a combination of computational approaches to distinguish variants with nonfunctional characteristics from disease-associated mutations with more dangerous consequences, and a credibility tool to enable ALS researchers to objectively assess the evidence for gene causation in ALS. Furthermore, integration of external tools, systems for feedback, annotation by users, and two-way links to collaborators hosting complementary databases further enhance the functionality of ALSoD.

Hum Mutat 33:1345–1351, 2012.

KEY WORDS: ALSoD; amyotrophic lateral sclerosis; Web–bases; database; genetics; motor neuron disease

Introduction

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease or Charcot disease, is the most common form of adult onset motor neuron degeneration, with a peak age of onset in the sixth or seventh decades [Valdmanis et al., 2009; Van Damme and Robberecht, 2009]. The prognosis is very poor with median survival of about three years from symptom onset. The causes of ALS are gradually being identified, but despite extensive research and rapidly increasing knowledge of the disease mechanisms, there remains no cure [van Es et al., 2010]. A disease-modifying drug treatment exists in the form of riluzole, a benzothiazole derivative that has a modest effect on survival. Noninvasive ventilation also extends survival and improves quality of life [Wijesekera and Leigh, 2009].

Although ALS is generally considered a single disease entity, there are various classifications based on genetic and phenotypic patterns, and it is probably more appropriate to consider it a syndrome of motor neuron degeneration with multiple causes. Advances in technology mean there is an enormous increase in the volume of research data produced, and a corresponding need for storage, analysis, and interpretation, particularly as our understanding of the relationships between genotype and phenotype mature.

An effective method for the curation and organization of such data is the online database [Lill et al., 2011]. Mutation databases of human genes are now highly prominent in all areas of healthcare, and experts in genetic diseases may curate published and unpublished mutations in locus-specific databases (LSDB) [Claustres et al., 2002; Fokkema et al., 2011]. These databases of Mendelian disease mutations play a fundamental role in research, diagnostic, and genetic healthcare [George et al., 2008]. Although the original role of databases was the simple storage of data, modern databases often have a related computational, bioinformatics, or analytical role that encourages interpretation of data.

Here, we present such a system in the form of the ALS Online Database (ALSoD) as seen in Figure 1: Homepage of ALSoD (http://alsod.iop.kcl.ac.uk). This freely available database has been transformed from a single gene storage facility recording mutations in the *SOD1* gene to a multigene ALS bioinformatics repository and analytical instrument combining genotype, phenotype, and geographical information with associated analysis tools.

ALSoD is not only a central repository for storing genetic information on the more than 100 ALS-related genes reported to date, but also shows graphs of the gender, age of onset, phenotype, and family history distributions of patient data stored on the database broken down by gene, mutation, or phenotypic group. The various analytical devices include a comparison tool to evaluate genes side by side or jointly with user configurable features, a pathogenicity prediction tool using a combination of computational approaches to distinguish variants with nonfunctional characteristics from disease-associated mutations that have more dangerous

*Correspondence to: Ammar Al-Chalabi, Department of Clinical Neuroscience, P041, Institute of Psychiatry, King's College London, London SE5 8AF, UK. E-mail: ammar.al-chalabi@kcl.ac.uk

Contract grant sponsors: ALS Association (1633); European Community's Health Seventh Framework Programme FP7/2007-2013 (grant agreement 259867); National Institute for Health Research (NIHR) Dementia Biomedical Research Unit at South London and Maudsley NHS Foundation Trust and King's College London (to A.A.C.).

Figure 1. The homepage of ALSoD (http://alsod.iop.kcl.ac.uk).

consequences, and a credibility tool to enable ALS researchers to objectively assess the evidence for association with ALS. Furthermore, integration of external tools, systems for feedback and annotation by users, and two-way links to collaborators hosting complementary databases further enhance the functionality of ALSoD.

Database Structure and Overview

Funding and Sponsorship

The ALSoD database is a joint project of the World Federation of Neurology and European Network for the Cure of ALS, and is funded through grants from the ALS Association, Motor Neurone Disease Association, ALS Canada, ALS Therapy Alliance, and MNDA Iceland.

Programming

Open source programming software such as JavaScript, C#, T-SQL, Perl, XML, and VB.NET integrated under the ASP.NET platform are implemented to write codes and scripts. ALSoD uses the Microsoft .NET framework. Microsoft SQL server 2008 is used to manage the database stored on the VM3 server of the Institute of Psychiatry, King's College London. Microsoft Visual Web Developer 2008 Express Edition is used to develop the user-interface dynamic Web pages. Google AJAX search and Google Earth API have been combined to overlay geographical mutation information of ALS-related genes on the globe.

Schema

The database schema now allows for flexibility and expansion because of changed table designs, rewritten queries, and implementation of stored procedures. Redundant tables have been removed, and a more supple structure is in place. The original database has been archived. New genes have been added to the tables, and a facility to easily add further genes designed. ALSoD now permits only registered users to submit novel gene, mutation, and patient data, and this is regularly validated by an ALS expert.

Web Design

A facelift was given to the Web page and some pages were redesigned for better visual representation of data. The Graphical User Interface allows data to be interpreted and viewed at a glance instead of using the tabular format of viewing data.

Collaborations and Embedded Tools

ALSoD uses third party open source bioinformatics tools to embed computational analysis within the database using Java applets. For example, in Figure 2, a screenshot of the Multiple Alignment and Mutations on *SOD1* gene using a combination of Claustalw and Jalview [Waterhouse et al., 2009] is used to provide multiple sequence alignments in other species for selected genes. GeneMANIA [Warde-Farley et al., 2010] allows users to select genes of interest for prediction of interactions. A Google Earth API is used for viewing maps of mutation, risk, and exposure distributions. Because many ALS gene variants are found in both familial and apparently sporadic ALS, a two-way link out to the ALSGene database provides evidence

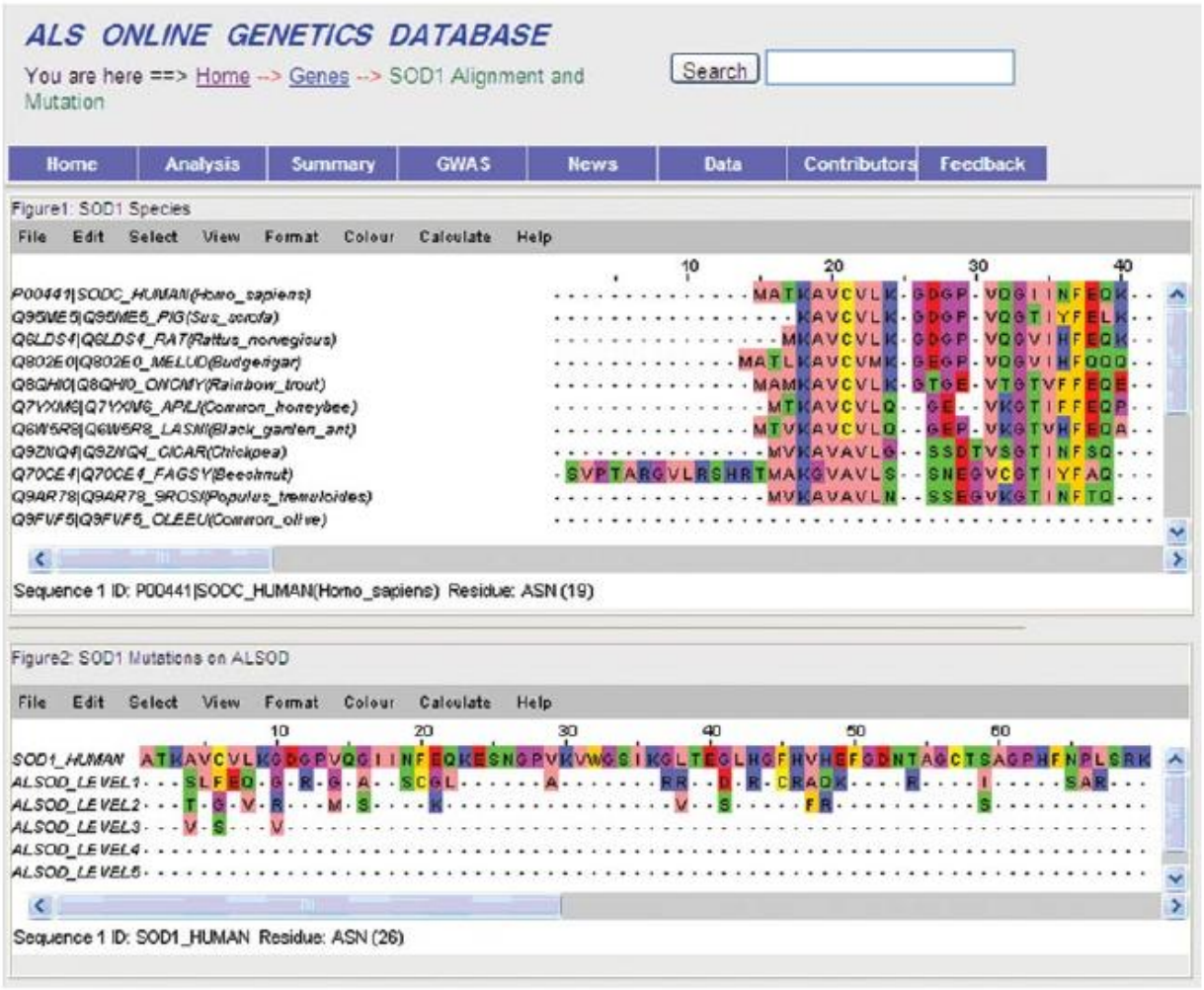


Figure 2. A screenshot of the Multiple Alignment and Mutations view of the *SOD1* gene using a combination of Claustalw and Jalview.

of association to complement the genotype–phenotype correlation available from familial ALS information in ALSoD [Lill et al., 2011]. A similar link out to fALS Connect, which is a collaboration between multiple interested agencies in the United States, including the patient organization The ALS Association and the research group The Northeast ALS (NEALS) Clinical Trials Consortium, makes ALSoD relevant for patients and carers as well as the scientific community. The database is adopted into the Human Variome Project (http://www.humanvariomeproject.org) and the GWAS Phenomap Project (http://www.gwascentral.org/gwasphenomap).

Integrated Bioinformatics Links

To avoid bias, users can retrieve gene-specific information through external links which have been programmed automatically for each gene, and which open in new windows. Unique identifiers are utilized by systematically linking to broad databases and bioinformatics tools freely available online. The scientific and nonscientific external links integrated into ALSoD include HGNC [White et al., 1997], Entrez Gene [Maglott et al.], UCSC Browser [Fujita et al.], Protein Structure [Rose et al.], OMIM [Amberger et al., 2011; Amberger et al., 2009], Genecards [Safran et al.], ProtScale [Gasteiger et al., 2005], KEGG [Kanehisa et al., 2000], Uniprot [Jain et al., 2009], iHop [Hoffmann and Valencia, 2004], Pathway in KEGG [Kanehisa et al., 2000], GeneTest [Pagon et al., 2002], AmiGO [Carbon et al., 2009], Ensembl [Hubbard et al., 2009], NCBI [Sherry et al., 2001], Life Science DB (Japan) [Yoshida et al., 2010], ALSGene [Lill et al., 2010], GeneWiki [Huss III et al., 2008], WolframAlpha (Maret), and WikiGenes [Hoffmann, 2008].

Feedback

Feedback is gained in two main ways: a Facebook page for ALSoD http://www.facebook.com/srch.php#!/pages/ALSoD/307667685943735, and a direct feedback page on the ALSoD Website. Comments are publicly displayed and a reCAPTCHA tool displays texts readable only by human users to prevent spammers from infiltrating the system. A news page generates automated summaries of ALS genetics news; and surveys conducted through the freely available online survey tool "SurveyMonkey" are embedded in the user interface.

In addition, by tracking the registered country of origin of page viewing and download requests, accessibility of the ALSoD database to the international ALS community can be monitored directly.

Changelog

ALSOD v0.1 Beta

First online in 1995 ALSOD (as formerly known) which was hosted at www.alsod.org was developed to store genetic and clinical information and to assist researchers in identifying correlations between phenotype and genotype in ALS for SOD1 mutations. The data available in the database were purely for the *SOD1* gene as this was the only available familial gene linked to ALS at the time [Radunovic and Leigh, 1999].

ALSOD v1.0

In 1999, the database was first fully functional and available for the research community.

ALSOD v2.0

In 2008, about 100 different mutation points across the *SOD1* sequence with corresponding clinical information were collated. Genetic mutations of the *SOD1* protein were linked to the hypothetical 3D mutant structure hosted on a University College London server developed by Andrew Martin's team [Wroe et al., 2008]. Fifty users from 17 institutions registered with ALSOD to submit ALS patient and mutation data. Ninety-seven familial individuals with 122 mutation data on *SOD1* were stored.

The website was relocated to http://alsod.iop.kcl.ac.uk/als following loss of the alsod.org domain. Data could be downloaded freely and the database queried to look for a specific mutation type in four ALS genes (*SOD1*, *ALS2*, *VAPB*, *NEFH*) or for specific information on patient data.

ALSoD v 3.0 Current Structure

ALSoD is now a relational database with a massive increase in available data through submissions by researchers and regular update by the database curators. The schema has been redesigned for uncomplicated future addition of familial and sporadic ALS patient data, associated mutations, and published ALS genes.

More than 100 ALS-related genes have now been added to the database with a current total of 431 mutations (195 pathogenic) and 589 patient data. Fifteen of the mutations are unpublished except in ALSoD. ALSoD Web pages have been visited over 280,000 times since 2009 by more than 22,900 unique visitors from 140 countries. There are 26 registered contributors excluding those from the host institution. Thirty-three different publications have cited functionalities or updates available on ALSoD.

Database Functions

Examples of Functionality

Users are able to summarize the relationship between mutational data and clinical patient data visually. On every gene overview page the total number of mutations and patients collated from publications is displayed, as well as associated phenotypic information such as the limb to bulbar ratio, age of onset as a box plot, male to female ratio, and familial to sporadic ratio. Key publications for each gene are listed and can be sorted by name of first author, year, or title. A section for genetic variations with their base pair positions, associated statistical results, author, year, and title of publication (where available) are shown in tabular format.

The data stored on ALSoD are also available through statistical reports (Fig. 3) via the reports tool at http://alsod.iop.kcl.ac.uk/Statistics/report.aspx. For example, at the moment this shows the average age of onset for ALS for all deposited genes is 45 years, with a range from 1 (*ALS2*) to 67 years old (*CRYM*). The top 20 most frequent mutations are shown with links to more details on each. The first six are *SETX*: Leu389Ser recorded in 38 out of 54 patients, *FUS*: Arg521His, recorded in 33 out of 72 patients, *SOD1*: Leu144Phe, recorded in 26 out of 194 patients, *UBQLN2*: Pro497His, recorded in 19 out of 35 patients, *VAPB*: Pro56Ser, recorded in 18 out of 19 patients, and *TARDBP*: Ala382Thr, recorded in 11 out of 75 patients.

If a mutation is subsequently reported in another individual, this can be seen at http://alsod.iop.kcl.ac.uk/Statistics/pathogenicity.aspx. This is important as it provides stronger evidence of pathogenicity. For example, the Arg521Cys mutation in the *FUS* gene originally reported from France has more recently been found in Italy. There is currently no system for flagging mutations subsequently found in controls.

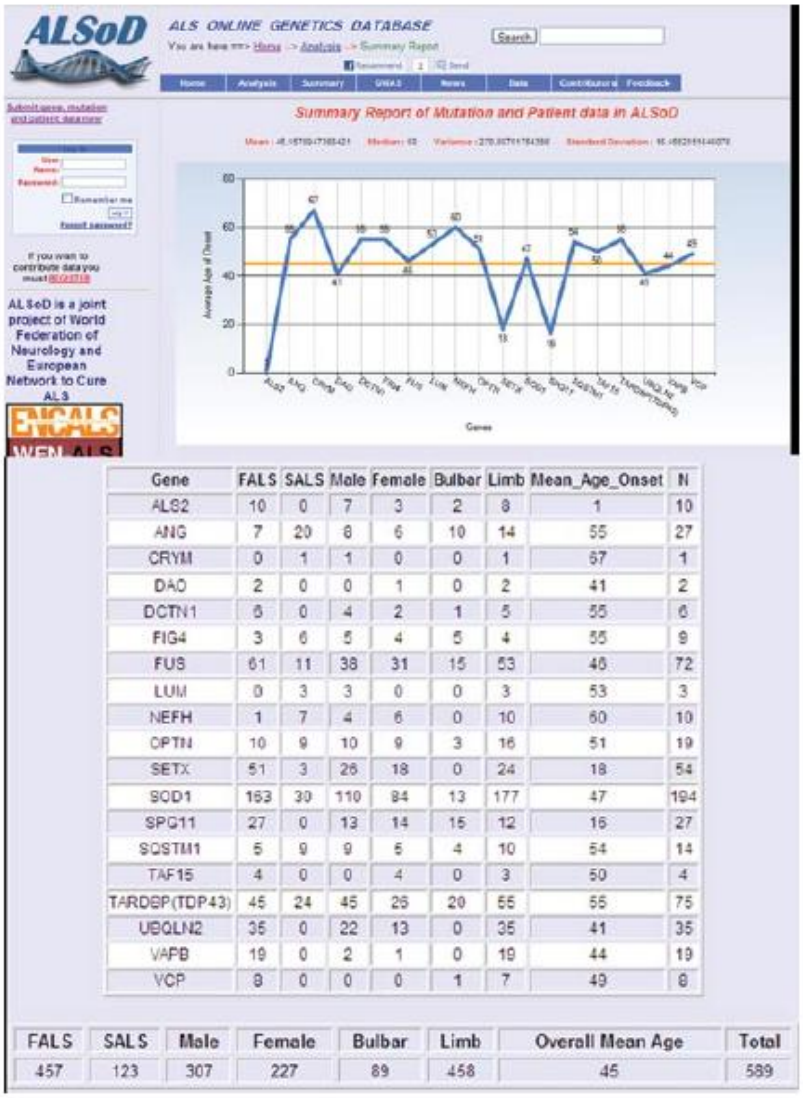
Where there are sufficient data that genotype–phenotype correlations are possible, a comparative study of selected genes can be performed on the detailed analysis page (Fig. 4) at http://alsod.iop.kcl.ac.uk/Statistics/analysis.aspx. A user-configurable form appears for users to choose two genes to compare, and if needed, the query can be restricted to patients within a particular age of onset range. For example, we can compare available data for *TARDBP* and *FUS*. The analysis shows that 73% of patients with *TARDBP* mutation have limb onset, whereas for *FUS* this is 78%; 63% of those with *TARDBP* mutation are male, compared with 55% for *FUS*, and there is a family history of ALS in 65% of those with *TARDBP* mutation compared with 85% of those with *FUS* mutation. The mean age of onset for patients with *TARDBP* mutation is higher at 55 than those with *FUS* mutation at 46 years. Geographical data are also deposited, showing in this case that most recorded *TARDBP* mutations have come from Italy (32%) but 11 countries have reported cases, with mutations found in Europe, North America, and Asia. For *FUS*, the situation is quite different, with just five countries reporting mutations, most of them from Belgium (43%).

A more detailed tabular format of the charts displayed is shown at the bottom of the analysis page including the author generating the data, linked to the PubMed abstract using pubmed ID. This is particularly useful for researchers who would like to verify the information presented.

Walk-Through

Scenario: A researcher is interested in understanding how *SOD1* mutations relate to ALS, and in particular, if there are any codons that are more likely to be mutated than others.

At the homepage the researcher can see that *SOD1* is listed as causative for ALS with a description of ALS1. Selecting this gene by clicking "Select" brings up an overview page. The header lists the gene, name and alternatives, the gene inheritance category (in this case familial ALS genes also found in sporadic ALS), the category of gene function (oxidative stress), the locus, two one-sentence



Gene	FALS	SALS	Male	Female	Bulbar	Limb	Mean_Age_Onset	N
ALS2	10	0	7	3	2	8	1	10
ANG	7	20	8	6	10	14	55	27
CRYM	0	1	1	0	0	1	57	1
DAO	2	0	0	1	0	2	41	2
DCTN1	6	0	4	2	1	5	55	6
FIG4	3	6	5	4	5	4	55	9
FUS	61	11	38	31	15	53	46	72
LUM	0	3	3	0	0	3	53	3
NEFH	1	7	4	6	0	10	60	10
OPTN	10	9	10	9	3	16	51	19
SETX	51	3	26	18	0	24	18	54
SOD1	163	30	110	84	13	177	47	194
SPG11	27	0	13	14	15	12	16	27
SQSTM1	5	9	9	5	4	10	54	14
TAF15	4	0	0	4	0	3	50	4
TARDBP(TDP43)	45	24	45	26	20	55	55	75
UBQLN2	35	0	22	13	0	35	41	35
VAPB	19	0	2	1	0	19	44	19
VCP	8	0	0	0	1	7	49	8

FALS	SALS	Male	Female	Bulbar	Limb	Overall Mean Age	Total
457	123	307	227	89	458	45	589

Figure 3. A summary report of mutation and patient data in ALSoD.

summaries, and the total number of mutations recorded with the total number of patients with genotype data. In this case, there are 165 mutations in 264 individuals. Below this is a graphic of the clinical presentation showing 93.4% of those with phenotype information had limb onset, a pair of graphs showing the distribution of age of onset, and two pie charts showing the male–female ratio and the proportion of familial and sporadic cases. The raw numbers are given in a table below the graphic. There are then a number of bioinformatics links and a literature review. Finally, a list of references is given for every deposited mutation.

Clicking the diagram link takes the researcher to a diagram of the first 106 pathogenic nonsynonymous mutations. The Google Map link presents a Google Earth display with the country of origin of each mutation highlighted.

The researcher is particularly interested in the A4V mutation. Using the analysis tools to predict pathogenicity, the researcher finds that A4V is predicted to be pathogenic, and the graphics below show that it has only been reported in familial ALS, with a mean age of onset of 47 and origin in the USA, Sweden, and Canada.

Comparison of the *SOD1* gene with *TARDBP* shows that bulbar onset is much commoner for *TARDBP* at 27%, the mean age of onset is higher and the proportion with no family history of ALS is also higher. Both genes are found mutated in many countries.

Analysis of the interaction networks of the two genes (under Analysis, Interactions) shows there are only a few links between the networks. Adding in *FUS* shows it to be a close interaction partner of *TARDBP* but not of *SOD1*.

Discussion

The ALSoD database collects genotype and phenotype information on ALS genes directly deposited by researchers or reported in publications, and is one of the oldest such databases in continuous use. Through links and collaborations with other databases such as ALSGene, and through analysis tools built into ALSoD, it is able to reveal patterns in data that would not otherwise be visible, and acts as a continuous review of ALS genetics and the correlation of genotype with phenotype.

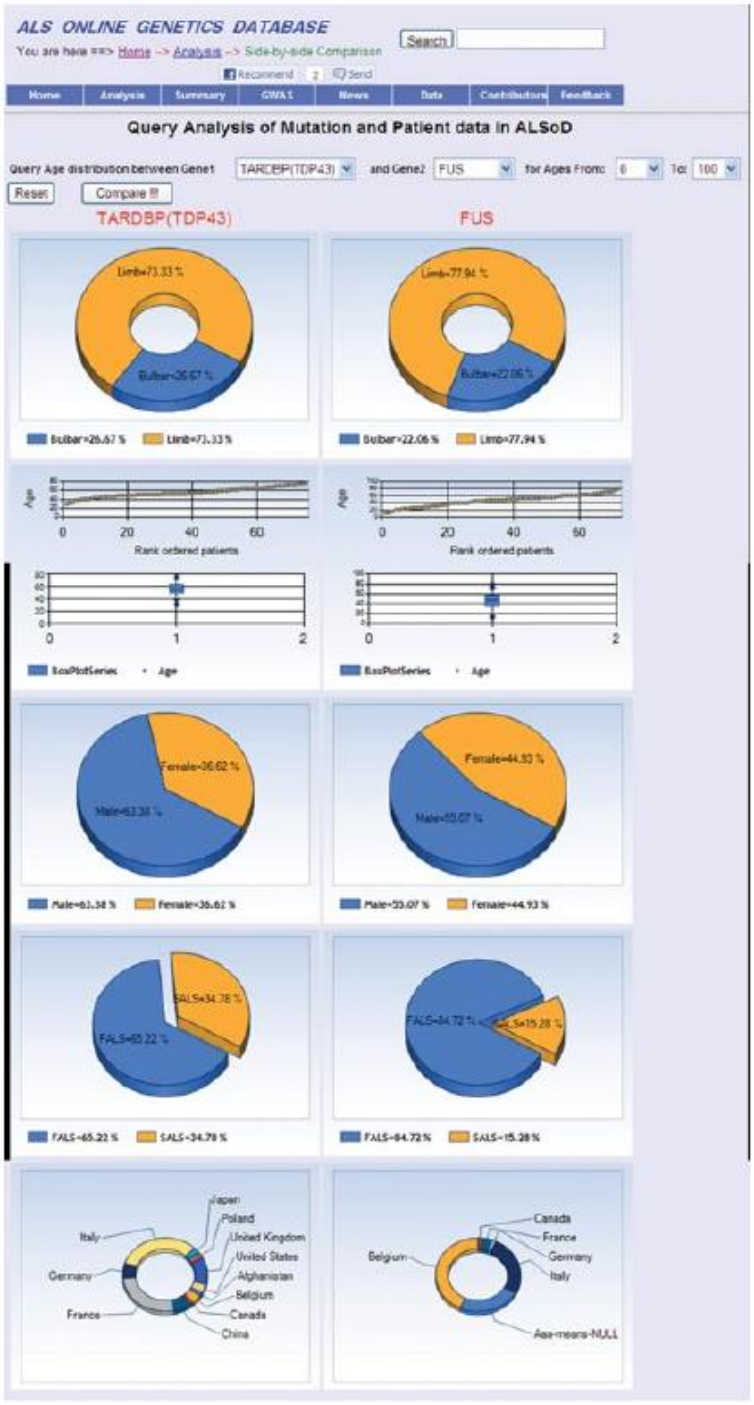


Figure 4. An example of side-by-side comparison of two genes.

A few issues with the original design meant that registration and access were cumbersome, discouraging users from registering. In addition, the database relied completely on the goodwill of the research community for updates, there was no easy way to include new genes, and there was no easy way to incorporate the latest advances in bioinformatics. These issues have all been addressed in modernizing ALSoD over the last few years.

In the field of ALS, as in other areas, available genetic data, phenotypic classifications, and relationships between mutations and clinical presentation have grown rapidly, and it can be difficult for those not involved in genetics on a daily basis to keep up. ALSoD fulfills an essential role in integrating and collating this otherwise overwhelming dataset into an understandable and manageable form.

A further difficulty faced by anyone interested in ALS genetics is that the number of putatively associated genes is far larger than the number of genes most people would accept as involved. ALSoD helps with this indirectly because the amount of available data on a particular gene corresponds with the amount of research and, therefore, to some extent, with the credibility of that gene as an ALS gene. We plan to formalize this in the future with an algorithm designed to objectively rank genes by the level of evidence supporting their role in ALS. This approach is already taken by databases such as ALSGene in association studies [Lill et al., 2011], and implementing it for familial ALS genes and for genotype–phenotype correlations is a logical extension of this process. More bioinformatics hyperlinks and integration with websites for the use of clinicians and nonclinicians, scientists, and the lay public are also under development [Wroe et al., 2008]. These include a section to deal with high throughput sequencing data. Data supporting a causative gene variation can be dealt with in the existing infrastructure, but what is more difficult is managing sequence data without clear annotation on the likely relationship of any variants seen with ALS. We will develop a tool for storage and display of such data to facilitate the identification of common patterns once sufficient sequences are deposited.

Feedback from users has suggested that extension of ALSoD to include integrated information from other species such as mouse and drosophila would be welcomed. We will begin the process of linking to the relevant external databases and collating published evidence. Once we have a critical mass of information, we will seek collaboration with experts in each field to help with analysis and presentation of data.

Having grown from a single gene mutational database, the ALS Online Database is now a major repository for all ALS genes with multiple tools for researchers familiar with ALS genetics and summaries for those wishing to keep up with genetic advances. With the support and feedback of the ALS research community, it will continue to develop and expand providing a continuous review of ALS genetics.

Acknowledgments

We are especially grateful for the long-standing and continued funding of this project from the ALS Association and the MND Association of Great Britain and Northern Ireland. We also thank ALS Canada, MNDA Iceland, and the ALS Therapy Alliance for support. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. Aleks Radunovic, Nigel Leigh, and Ian Gowrie originally conceived ALSoD. ALSoD is a joint project of the World Federation of Neurology (WFN) and European Network for the Cure ALS (ENCALS).

Conflicts of Interest: A.A.C. is on the Scientific Advisory Board of the ALSGene database, the Biomedical Research Advisory Panel of the Motor Neurone Disease Association, and is a consultant for Biogen Idec and Cytokinetics. He receives royalties for the books "The Brain: A Beginner's Guide" (Oneworld Publications) and "The Genetics of Complex Human Diseases" (Cold Spring Harbor Laboratory Press).

References

Amberger J, Bocchini C, Hamosh A. 2011. A new face and new challenges for Online Mendelian Inheritance in Man (OMIM®). Hum Mutat.

Amberger J, Bocchini CA, Scott AF, Hamosh A. 2009. McKusick's Online Mendelian Inheritance in Man (OMIM®). Nucleic Acids Res 37(suppl 1):D793–796.

Carbon S, Ireland A, Mungall CJ, Shu SQ, Marshall B, Lewis S. 2009. AmiGO: online access to ontology and annotation data. Bioinformatics 25:288–289.

Claustres M, Horaitis O, Vanevski M, Cotton R. 2002. Time for a unified system of mutation description and reporting: a review of locus-specific mutation databases. Genome Res 12:680–688.

Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. 2011. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 32:557–563.

Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A. The UCSC Genome Browser database: update 2011. Nucleic Acids Res 39(suppl 1):D876–882.

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook 571–607.

George R, Smith T, Callaghan S, Hardman L, Pierides C, Horaitis O, Wouters M, Cotton R. 2008. General mutation databases: analysis and review. J Med Genet 45:65–70.

Hoffmann R. 2008. A wiki for the life sciences where authorship matters. Nat Genet 40:1047–1051.

Hoffmann R, Valencia A. 2004. A gene network for navigating the literature. Nat Genet 36:664–664.

Hubbard T, Aken B, Ayling S, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Clarke L. 2009. Ensembl 2009. Nucleic Acids Res 37(suppl 1):D690–697.

Huss III JW, Orozco C, Goodale J, Wu C, Batalov S, Vickers TJ, Valafar F, Su AI. 2008. A gene wiki for community annotation of gene function. PLoS Biol 6:e175.

Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, Martin MJ, McGarvey P, Gasteiger E. 2009. Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinformatics 10:136.

Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. 2000. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res 38(suppl 1):D355–360.

Lill C, Zauft U, Roehr J, Meissner E, Schjeide B, Schjeide L, McQueen M, Bertram L. 2010. The ALSGene Database: systematic meta-analyses and field synopsis of genetic association studies in ALS. In: 21st International Symposium on ALS/MND. p Abstract #SW264, Florida.

Lill CM, Abel O, Bertram L, Al-Chalabi A. 2011. Keeping up with genetic discoveries in amyotrophic lateral sclerosis: the ALSoD and ALSGene databases. Amyotroph Lateral Scler 12:238–249.

Maglott D, Ostell J, Pruitt KD, Tatusova T. 2011. Entrez Gene: gene-centered information at NCBI. Nucleic Acids Res 39(suppl 1):D52–57.

Maret S. Wolfram| Alpha Knowledgebase. Charleston Advisor 12:48–50.

Pagon RA, Tarczy Hornoch P, Baskin PK, Edwards JE, Covington ML, Espeseth M, Beahler C, Bird TD, Popovich B, Nesbitt C. 2002. GeneTests GeneClinics: genetic testing information for a growing audience. Hum Mutat 19:501–509.

Radunovic A, Leigh P. 1999. ALSODatabase: Database of SOD1 (and other) gene mutations in ALS on the Internet. Amyotroph Lateral Scler 1:45–49.

Rose PW, Beran B, Bi C, Bluhm WF, Dimitropoulos D, Goodsell DS, Prli A, Quesada M, Quinn GB, Westbrook JD. 2011 The RCSB Protein Data Bank: redesigned web site and web services. Nucleic Acids Res 39(suppl 1):D392–401.

Safran M, Dalah I, Alexander J, Rosen N, Stein TI, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H. GeneCards Version 3: the human gene integrator. Database: J Biol Databases Curation 2010:baq20.

Sherry S, Ward M, Kholodov M, Baker J, Phan L, Smigielski E, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29:308–311.

Valdmanis PN, Daoud H, Dion PA, Rouleau GA. 2009. Recent advances in the genetics of amyotrophic lateral sclerosis. Curr Neurol Neurosci Rep 9:198–205.

Van Damme P, Robberecht W. 2009. Recent advances in motor neuron disease. Curr Opin Neurol 22:486–492.

van Es MA, Dahlberg C, Birve A, Veldink JH, van den Berg LH, Andersen PM. 2010. Large-scale SOD1 mutation screening provides evidence for genetic heterogeneity in amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 81:562–566.

Warde-Farley D, Donaldson S, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes C. 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res 38(Web Server issue):W214–220.

Waterhouse AM, Procter JB, Martin D, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25:1189–1191.

White J, McAlpine P, Antonarakis S, Cann H, Eppig J, Frazer K, Frezal J, Lancet D, Nahmias J, Pearson P. 1997. Guidelines for human gene nomenclature (1997). HUGO Nomenclature Committee. Genomics 45:468–471.

Wijesekera LC, Leigh PN. 2009. Amyotrophic lateral sclerosis. Orphanet J Rare Dis 4:3.

Wroe R, Wai-Ling Butler A, Andersen P, Powell J, Al-Chalabi A. 2008. ALSOD: the amyotrophic lateral sclerosis online database. Amyotroph Lateral Scler 9:249–250.

Yoshida M, Takahashi Y, Koike A, Fukuda Y, Goto J, Tsuji S. 2010. A mutation database for amyotrophic lateral sclerosis. Hum Mutat 31:1003–1010.

Chapter 3 Publication 2 - Keeping up with genetic discoveries in amyotrophic lateral sclerosis: The ALSoD and ALSGene databases [486]

Contribution: Joint First Author

Journal: Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration

Volume: July 2011, Vol. 12, No. 4 , Pages 238-249

Authors: Christina M. Lill, Olubunmi Abel, Lars Bertram & Ammar Al-Chalabi

Pubmed ID: 21702733

DOI: 10.3109/17482968.2011.584629

Link: http://informahealthcare.com/doi/abs/10.3109/17482968.2011.584629

Amyotrophic Lateral Sclerosis, 2011; 12: 238–249

informa
healthcare

ORIGINAL ARTICLE

Keeping up with genetic discoveries in amyotrophic lateral sclerosis: The ALSoD and ALSGene databases

CHRISTINA M. LILL[1,2], OLUBUNMI ABEL[3], LARS BERTRAM[1] & AMMAR AL-CHALABI[3]

[1]*Neuropsychiatric Genetics Group, Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin, Germany,* [2]*Department of Neurology, Johannes Gutenberg University, Mainz, Germany, and* [3]*MRC Centre for Neurodegeneration Research, King's College London Institute of Psychiatry, London, UK*

Abstract

Amyotrophic lateral sclerosis (ALS) is a genetically heterogeneous disorder that shows a characteristic dichotomy of familial forms typically displaying Mendelian inheritance patterns, and sporadic ALS showing no or less obvious familial aggregation. While the former is caused by rare, highly penetrant, and pathogenic mutations, risk for sporadic ALS is probably the result of the combined effects of common polymorphisms with minor to moderate effect sizes. Owing to recent advances in high-throughput genotyping and sequencing technologies, genetic research in both fields is evolving at a rapidly increasing pace making it more and more difficult to follow and evaluate the most significant progress in the field. To alleviate this problem, our groups have created dedicated and freely available online databases, ALSoD (http://alsod.iop.kcl.ac.uk/) and ALSGene (http://www.alsgene.org), which provide systematic and in-depth qualitative and quantitative overviews of genetic research in both familial and sporadic ALS. This review briefly introduces the background and main features of both databases and provides an overview of the currently most compelling genetic findings in ALS derived from analyses using these resources.

Key words: *ALSoD, ALSGene, amyotrophic lateral sclerosis, ALS, database, genetic, UNC13A*

Introduction

For amyotrophic lateral sclerosis (ALS), as well as for many other neurodegenerative disorders such as Alzheimer's disease (1) and Parkinson's disease (2), familial aggregation was recognized as a salient feature decades before any of the underlying biological and biochemical properties were known (3). In many cases, the identification of specific, disease-segregating gene mutations has first directed the attention of molecular biologists to certain proteins and pathways. The breakthrough in understanding ALS pathogenesis started with the discovery of autosomal-dominant mutations in *SOD1* on chromosome 21q22 (encoding soluble copper/zinc superoxide dismutase 1) in 1993 (4). In the last decade, 12 additional genes causing various forms of ALS have been identified (Table I), and it can be anticipated that the number of ALS-causing mutations will increase further with the application of recently developed massively parallel sequencing techniques.

Another similarity to other neurodegenerative diseases is the prominent dichotomy of these rare, familial forms of ALS following Mendelian inheritance (5–10% of all cases) alongside seemingly 'sporadic' forms of the disease showing no or less obvious familial aggregation, with no clear clinical or pathological distinction between the two. While a small percentage of the sporadic ALS cases have also been found to harbour mutations in the known Mendelian ALS genes, the aetiology for the majority remains largely unknown. It is likely that there is a contribution from common genetic variants (polymorphisms) that are also present in the healthy population (with frequencies > 1%). Each of these polymorphisms

Correspondence: The ALSGene database: L. Bertram, Max Planck Institute for Molecular Genetics, Ihnestr. 63, Berlin, Germany. E-mail: lbertram@molgen.mpg.de. The ALSoD database: A. Al-Chalabi, MRC Centre for Neurodegeneration Research, King's College London Institute of Psychiatry P 041, London SE5 8AF, UK. E-mail: ammar.al-chalabi@.kcl.ac.uk

(Received 10 March 2011; accepted 5 April 2011)

ISSN 1748-2968 print/ISSN 1471-180X online © 2011 Informa Healthcare
DOI: 10.3109/17482968.2011.584629

probably exerts only a small effect on ALS risk (odds ratios (ORs) in the order of ~1.2), so that it is via their combined action that they impact ALS susceptibility. Because of the small individual genetic effect sizes, thousands of subjects have to be tested to uncover these risk genes with sufficient confidence (i.e. to achieve genome-wide significance which is commonly defined as *p*-values falling at or below ~5×10^{-8}). Several hundred candidate-gene association studies have been performed in the field of ALS genetics during the last two decades, yielding mostly inconclusive results. In recent years, genome-wide association studies (GWAS), which can assess several million polymorphisms across the genome in one experiment, have suggested several new ALS risk genes, most of which are still awaiting independent follow-up and replication.

The growing amount of information in the field of ALS genetics is becoming increasingly difficult to follow, evaluate and interpret. This is true for both Mendelian and non-Mendelian, genetically complex ALS. To address this problem, the ALSoD and the ALSGene databases have been developed independently. Both databases aim to provide systematic and in-depth overviews of the respective ALS genetics fields. The aim of this review is to introduce both resources by outlining their complementary goals and methods, as well as by summarizing the most interesting recent findings of each project.

Assessing Mendelian mutations and rare variants in ALS: the ALS Online Genetic Database (ALSoD)

Although the first human disease-causing mutation was identified in haemoglobin in 1949, the pace of discovery was such that the relationship between gene variations and the diseases they caused could be easily remembered. As the numbers grew, it became important to store such information systematically, and so the first genetics database came about in the1970s when McKusick created a catalogue of mutations and inherited diseases, known as "Mendelian Inheritance in Man". A version of this database was later developed for online access (5), and is now known as OMIM. By 1994 there were regularly updated online lists of gene mutations as well as gene-specific databases. Key to the success of such resources was a universal and logical mutation naming convention based on amino acid changes (6,7).

Mutation databases of human genes are now becoming more prominent and important in all areas of health care. Published and unpublished mutations are typically curated in locus-specific databases (LSDB) (8), and apart from providing vital information on protein function and structure, databases of mutations causing Mendelian disease play a fundamental role in research, diagnostics, and genetic health care (9,10). The existence and regular upkeep of these databases is often vital to a broad range of research fields.

Table I. Genes proposed to cause Mendelian forms of ALS.

Gene	Protein	Location	Inheritance	#Mutations	#Patients	Proposed molecular effects/pathogenic relevance
ALS2	amyotrophic lateral sclerosis 2	2q33.1	recessive	19	4	endosome/membrane trafficking
ANG	angiogenin	14q11.1	dominant	18	27	RNAase
DAO	D-amino-acid oxidase	12q24	dominant	1	1	D-serine pathway
FIG4	FIG4 homolog, SAC1 lipid phosphatase domain containing (S. cerevisiae)	6q21	dominant	10	9	Golgi complex maintenance, vesicle trafficking, cell signalling
FUS	fused in sarcoma	16p11.2	both	27	5	RNA processing
NEFH	neurofilament, heavy polypeptide 200kDa, heavy chain	22q12.1-q13.1	?	7	10	axonal transport
OPTN	optineurin	10p13	recessive	3	0	Golgi complex maintenance and membrane trafficking
SETX	senataxin	9q34.13	dominant	4	60	DNA and RNA processing
SOD1	superoxide dismutase 1	21q22.11	both	156	184	toxic aggregation of SOD1; oxidative damage; mitochondrial dysfunction; RNA destabilization; impairment of axonal transport; glutamate excitotoxicity
TARDBP	TAR DNA binding protein (TDP-43)	1p36.22	dominant	38	66	RNA processing
VAPB	VAMP (vesicle-associated membrane protein)-associated protein B and C	20q13.33	dominant	2	19	vesicle trafficking
VCP	valosin-containing protein	9p13	dominant	4	8	ubiquination

For an up-to-date overview of these and other potential Mendelian ALS genes see the 'ALSoD database' (http://alsod.iop.kcl.ac.uk). Note that mutations in other genes have also been proposed to cause familial forms of ALS, albeit with inconclusive evidence (see text for more details).
Mutations = number of mutations reported as ALS-causing; # Patients = number of records on ALS patients submitted to ALSoD that carry mutations in this gene.

Following the identification of ALS-causing mutations in *SOD1* there was a great focus of research on this gene. As the number of mutations increased and the range of phenotypes expanded, the need arose for a database that might reveal new genotype-phenotype relationships, keep track of the different disease-causing genes, and aid the recording of gene variations that might not be published. From this came the ALS Online Database, ALSoD, an open source data repository designed for sharing clinical and genetic data, developed by a consortium within the World Federation of Neurology (11,12). It provides both the scientific community and wider public with up-to-date information on ALS-related genes identified from family studies, candidate gene studies or genome-wide association (Figure 1). The functional importance of such a database is underscored by the appearance of other similar websites, e.g. an ALS mutation database recently constructed as part of the Life Science Integrated Database Project conducted by the Japan Ministry of Education, Culture, Sports, Science, and Technology (13). This contains their original experimental results as well as published data extracted from scientific journals. This database is expected to play a complementary role to the ALSoD website, especially in collecting genetic variations commonly found in Asia. The ALS Online Database website was initially located at http://www.alsod.org and is now located at http://alsod.iop.kcl.ac.uk, but it remains a World Federation of Neurology resource, not tied to any institution. This change in web address was a simply the result of the need for huge data storage.

Figure 1. ALSoD homepage.

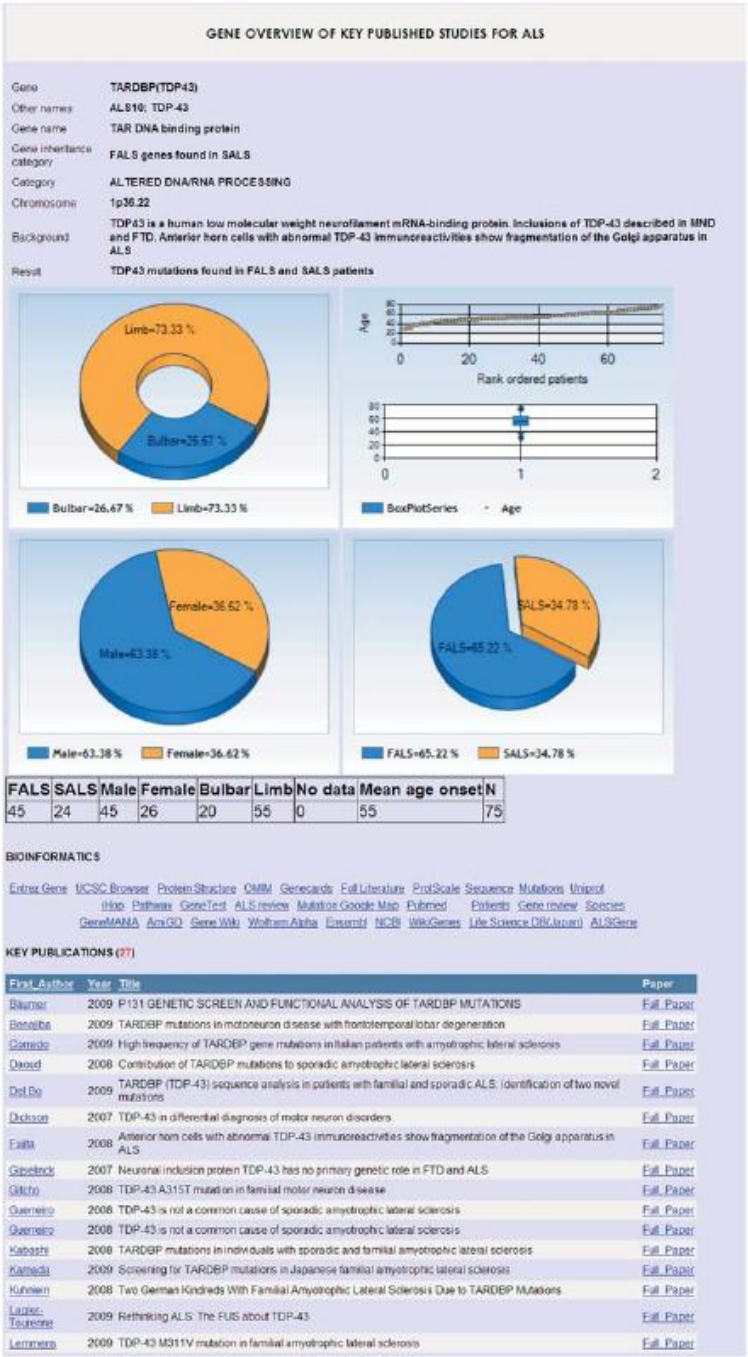


Figure 2. Genotype-phenotype summary for *TARDBP* (coding for TDP43). Gene summaries are displayed on selection. More detailed customised analyses are also possible allowing specific questions to be asked of the database.

Aims

ALSoD (freely available at http://alsod.iop.kcl.ac.uk) is designed to allow easy access to information about ALS-causing gene variants and their phenotypic effects. Because genotype-phenotype correlations are usually clearest for Mendelian disease genes, and the database has been in existence for several years, much of the focus is on familial ALS, but all currently known ALS-causing genes are represented. Where relevant there are links out to an appropriate database resource. For example, the evidence for association at a given gene is provided by a link out

Figure 3. An example of a geographical overview of the mutation distribution of *SOD1*. Geographical analysis is available for most FALS genes. More sophisticated map overlays are planned.

to the ALSGene database where meta-analysis and other explorations are possible (see below), and predictions of *SOD1* mutation effects on structure and function are provided by a link to a SOD1 protein structure database at UCL, whereas phenotype information on *TARDBP, FUS* or *OPTN* mutations is accessed locally. Mutation hot and cold spots, country-of-origin effects and literature and bioinformatics resources are also included within ALSoD.

Methods

ALSoD allows users to submit new gene, mutation and patient data. It also has various tables in the database schema, storing data on sequence mutation data on genes, and anonymized patient data including phenotypic features and country of origin. Published and unpublished data are continually curated by the database administrator and registered ALS researchers. Published data are obtained by regular PubMed searches, and screening of cross-references from relevant publications. Unpublished data are uploaded by users (as described above) or made available to the ALSoD investigators by collaborations, and are published on the ALSoD website after consistency checks performed by the database administrator. There are currently 12 contributors from various international institutions.

To allow integration of different databases through one gateway, ALSoD uses unique identifiers to systematically link to other bioinformatic tools and comprehensive databases. Selecting a link automatically interrogates the relevant third party website with the appropriate genetic information. Thus, it is possible to identify the most up-to-date list of genes implicated in ALS, explore their interactions and pathways, and review the relevant literature. Customized and routine analyses of mutational relationships with phenotype are possible, either for genes with phenotypic data, such as *SOD1*, *TARDBP*, *FUS*, *SETX*, and others, or as a general overview of familial and sporadic phenotypic patterns (Figure 2).

The mapping of gene variants on Google Earth allows researchers or patients to view the geographical distribution of reported gene variations associated with ALS (Figure 3). The same computational method allows us to map the origin of users and show which countries predominantly access ALSoD.

For users with their own association data, an on-the-fly analysis is available to combine the data available in ALSoD with unpublished user data that can be confidentially uploaded. The user data is formatted accordingly before upload and the result is fed back in minutes without storing users' data on the database.

Results

Currently, ALSoD displays information on 73 ALS genes (17 for familial ALS), 298 mutations and 419 patients (343 with age of onset data and described in analyses below). Typically there are more than 100 unique users visiting ALSoD each day, and in November 2010 the pages were accessed more than 18 000

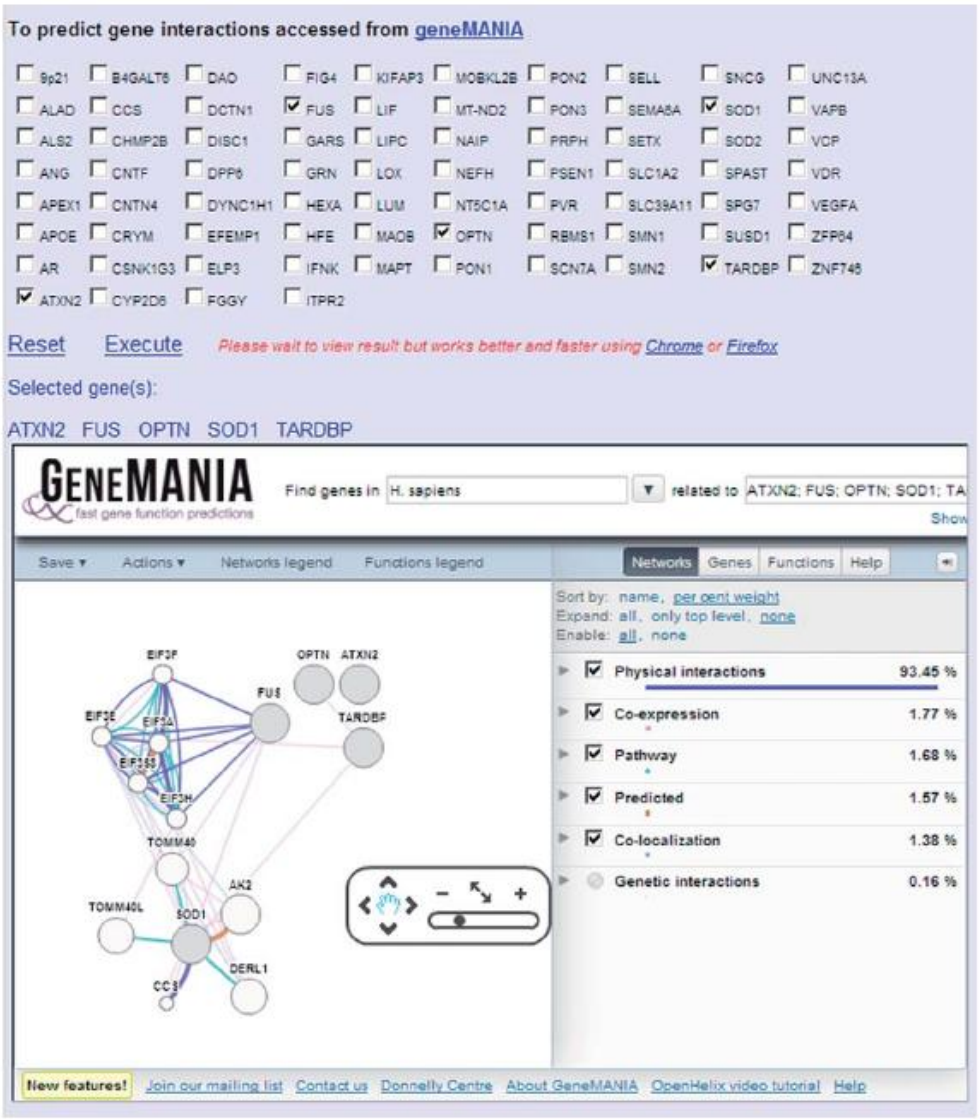


Figure 4. Gene interactions using a GeneMANIA interface at ALSoD. Potential gene pathways and networks can be explored via the GeneMANIA feature on ALSoD, e.g. by including established or proposed ALS genes in the functional predictions.

times. Data can be displayed or interrogated online or downloaded into a spreadsheet. For genes, the mutation type, sequence change, original and mutated amino acid, functional annotation and exon number are reported, with a link to the predicted mutant structure. For patients, the gene, mutation and mutation type are reported along with the patient's sex, country of residence, ethnicity, age at onset, clinical pattern, survival data and family history.

Analysis of the record of Mendelian mutations in which one would expect a 50:50 male: female ratio, shows that the actual recorded sex ratio is 172:119, which is almost exactly the 3:2 ratio commonly reported for ALS in general from clinic-based studies. While for 250 patients there is a family history of ALS, 84 patients have none despite a Mendelian mutation, and nine record it as unknown. Those without a family history are not skewing the sex ratio, however, as for those with a family history the sex ratio is 120:88, which is still very close to 3:2.

The mean age of onset overall is 45.7 years. Excluding those with juvenile onset (for example with *ALS2* gene mutation), mean age of onset is 46.2 years. The mean age of onset for those with a family history is 42.8 years while for those without it is 53.3 years, which is very close to the mean age of onset recorded in clinic based studies, and suggests that a family history is associated with a younger age of onset. However, for those with *SOD1* mutation, mean age of onset is 47.2 years ($n = 125$), which breaks down into 47.9 years for those with a

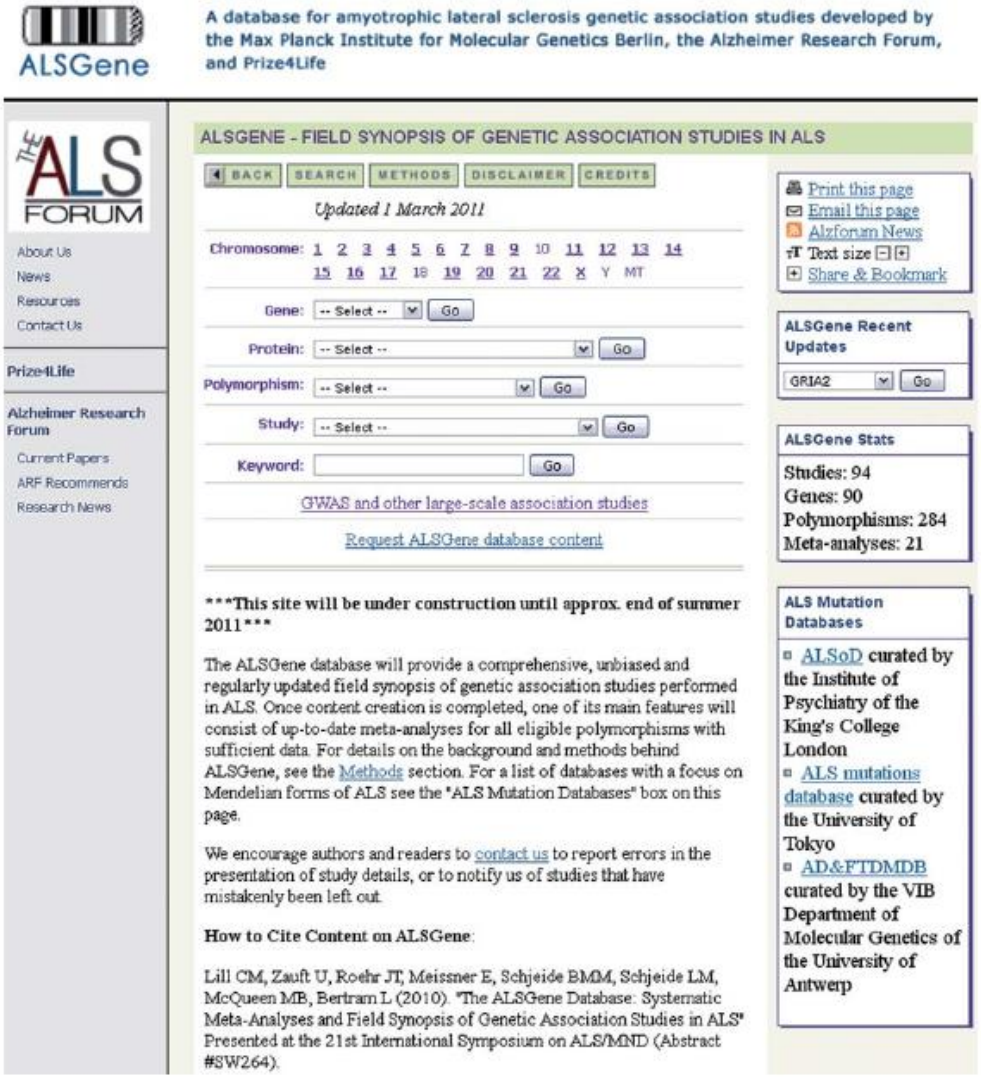


Figure 5. Screenshot of the ALSGene homepage (www.alsgene.org).

family history ($n = 105$) and, surprisingly, 41.7 for those without ($n = 19$). For *TARDBP* mutations (codes for TDP43 protein), mean age of onset is 55.5 years ($n = 75$) and for *FUS* mutations, 50.4 years ($n = 5$). There are 31 different *TARDBP* mutations recorded, the commonest being A382T ($n = 11$), G348C ($n = 8$), and M337V ($n = 7$). Five individuals with *SOD1* mutation have bulbar onset, whereas 116 have limb onset (30 upper limb, 83 lower limb, three not specified). Twenty individuals with *TARDBP* mutation have bulbar onset, compared with 41 limb onset (10 upper limb, four lower limb, rest not specified). This difference in recorded presentation patterns between those with *SOD1* and *TARDBP* mutations is highly significant (χ^2 p-value 4×10^{-7}).

A GeneMANIA web interface allows prediction of ALS gene interactions in ALSoD, which is a powerful tool for the exploration of new pathways (14) (Figure 4).

Future implementations

ALSoD will widen its phenotypic remit to include frontotemporal dementia genes, develop more detailed geographical displays (for example, overlaying epidemiology study data on Google Earth), integrate genome-wide sequence data with appropriate display tools and, through collaborations, integrate transcriptomics data with existing gene and phenotype data. LOD scores and sequence data will be ported to a genome browser track for UCSC and Ensembl browsers allowing for easy analysis together with existing data as well as comparison with ALSGene results. Through collaboration with ALSGene, association evidence for phenotypic modifiers for age of onset, site of onset and survival will b

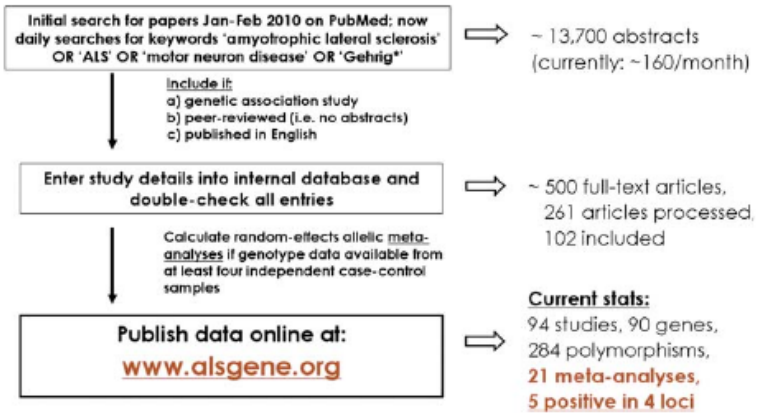


Figure 6. Summary of ALSGene methodology. Flowchart of data selection, processing and analysis strategies applied for ALSGene. Polymorphisms with non-overlapping data from at least four data sets are subjected to meta-analyses. Significant meta-analysis results are highlighted in a designated section on the ALSGene homepage.

Assessing genetically complex forms of ALS: the ALSGene database

The methods to identify susceptibility variants for genetically complex traits differ from those aimed at finding new causative disease genes. Most commonly, susceptibility variants are identified by probing for significant differences in allele or genotype frequencies of a genetic variant across groups of affected ('cases') versus unaffected ('controls') individuals by performing genetic association analyses. As outlined in the introduction, several hundred such genetic association studies have been published in the field of sporadic ALS over the past two decades. As a result, more than 150 putative ALS risk genes have been reported. These findings have been met with inconclusive replication evidence, at best. More often than not, presumed risk effects were not observed in independent datasets, complicating an easy interpretation of these findings. More recently, genome-wide approaches have been applied to case-control datasets (15–24) which – at least initially – did not lead to an improvement in replicability of the top findings. Concerted efforts using increasingly larger datasets have somewhat alleviated this situation (e.g. References 20,24). With a steadily growing number of GWAS and smaller-scale association studies focusing on sporadic ALS, the field becomes increasingly more difficult to evaluate. To facilitate this situation, our group has developed a systematic meta-analytic approach to genetic association studies that has already successfully been applied to other neurodegenerative diseases, e.g. Alzheimer's disease (25) and Parkinson's disease (26). We have begun applying the same systematic methodology to the field of sporadic ALS in the form of the ALSGene database.

Aims

ALSGene (freely available via the URL: http://www.alsgene.org or Prize4Life's ALS portal http://www.ResearchALS.org; see Figure 5 for a screenshot of the homepage of ALSGene) aims to serve as an exhaustive, unbiased, and regularly updated resource of genetic association studies in ALS. One of its key features will be up-to-date meta-analyses of all eligible genetic polymorphisms that have been investigated for association with ALS risk. This entails the inclusion of genome-wide association data, either in part or – if available – in full. All data and results are concisely summarized and displayed online. Upon completion of content curation, this will include an up-to-date list of 'Top Results', i.e. those showing the strongest evidence for association with risk for ALS when all available data are combined.

Methodology

For a visual summary of this paragraph see Figure 6. Eligible studies are identified via PubMed searches, or by screening the table of contents of neurological and genetic journals. Additional publications are also sought by systematically screening the references of relevant publications (e.g. primary association papers, but also reviews, etc.). To date, these searches have yielded ~500 articles on ALS genetics of which the full text is currently being screened for eligibility. A publication is included in ALSGene if it represents a genetic association study (here defined as studies investigating DNA sequence variants with $\geq 1\%$ frequency in the general population, a.k.a. 'polymorphisms'), and if it has been published in a peer-reviewed English language journal. Demographic details of included studies and genotype summary data of the investigated polymorphisms are extracted from each publication, and summarized on the ALSGene website. Each study included and listed in ALSGene is hyperlinked to the respective abstract in PubMed to facilitate identification and access to the original publication. In addition, an overview of all published ALS GWAS is provided in

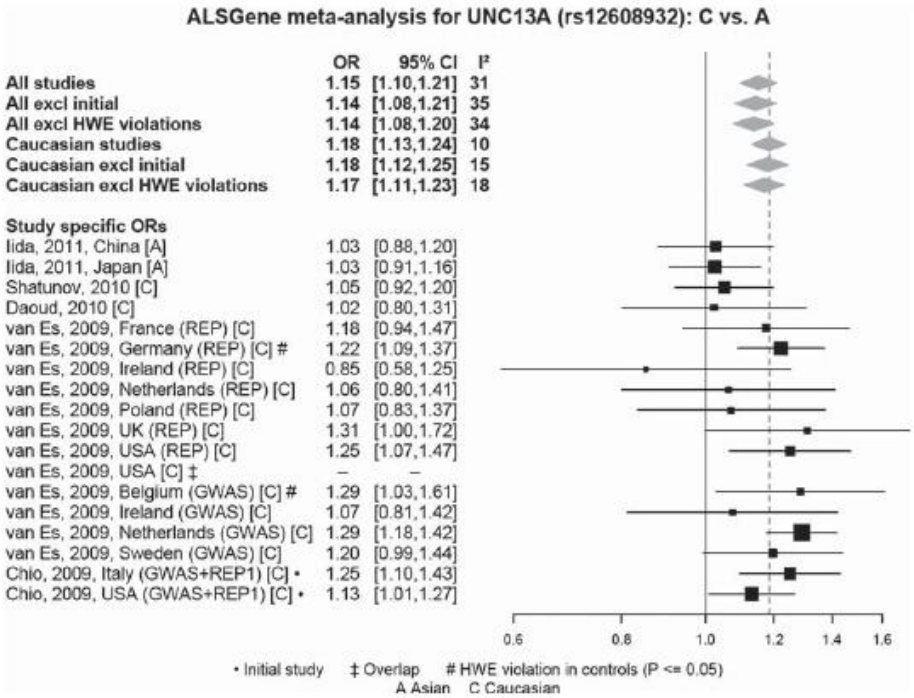


Figure 7. Random-effects meta-analysis of rs12608932 in *UNC13A* including all published studies (as of 1 March 2011). Example ALSGene forest plot providing a visual summary of meta-analysis results. Study-specific odds allelic ratios (ORs, bottom section) are displayed either as provided in the original publication or calculated de novo from allele or genotype summaries. Black squares represent point estimates of the OR (scaled to sample size), horizontal lines represent 95% confidence intervals. Summary ORs (top section) are then calculated from the study-specific ORs using random effects models for a number of different paradigms. Grey diamonds represent point estimates and 95% confidence intervals of the respective meta-analyses. I^2 = a measure of between-study heterogeneity (see text for more details). Forest plots are only available for polymorphisms that have data available from four independent data sets.

a dedicated section on ALSGene (http://www.alsgene.org/largescale.asp). Alongside the main characteristics of each study, this GWAS overview section also provides a hyperlinked list of 'featured genes', i.e. those loci or pathways highlighted by the primary authors as the main outcome of their study after having completed all analyses. ALSGene meta-analyses are performed for all genetic polymorphisms with at least four independent case-control datasets using allelic contrasts. Whenever possible, meta-analyses are complemented by GWAS data for overlapping polymorphisms (either directly genotyped or determined by imputation). Results of these meta-analyses are visualized on the website as forest plots (representing an up-to-date summary of the current evidence) and cumulative meta-analyses (summarizing the development of effect size estimates over time). Additional analyses include an assessment of the 'epidemiologic credibility' of significant findings using the HuGENet 'Venice' criteria (27,28), as well as Bayesian analyses (29). Furthermore, meta-analyses are carried out after stratification for distinct ethnicities (e.g. Asian and Caucasian ethnicities), when at least three independent datasets for the respective ethnicity are available (for an example see Figure 7). More details on the methods behind data collection and analyses for ALSGene can be found online. All significant meta-analysis results are summarized in a dedicated section of ALSGene ('Top Results'; note that the Top Results feature will only become available after primary literature-based content creation has been completed, i.e. approx. summer/fall of 2011).

Preliminary results

Currently, the online version of ALSGene displays the details and genotype summary data of over 90 independent studies that have investigated > 250 polymorphisms in 90 genes. As this paper is going to press, 21 meta-analyses are available, of which five show nominally significant results in four genetic loci (*APEX1* (OR 0.78; $p = 0.028$), *HFE* (OR 1.72; $p = 0.0034$), *UNC13A* (OR 1.18 in Caucasians; $p = 3.6 \times 10^{-13}$; Figure 7)), and a hitherto uncharacterized locus on chromosome 9p21.2 (OR 1.25; $p = 5.3 \times 10^{-18}$). The chromosome 9p21.2 locus and *UNC13A*, both initially implicated by recent GWAS (20,23,24), currently show association with risk for ALS at genome-wide significance (i.e. p-values $\leq 5 \times 10^{-8}$). The current *UNC13A* finding described here is a good example of why systematic data integration efforts are needed, even in the 'GWAS era'

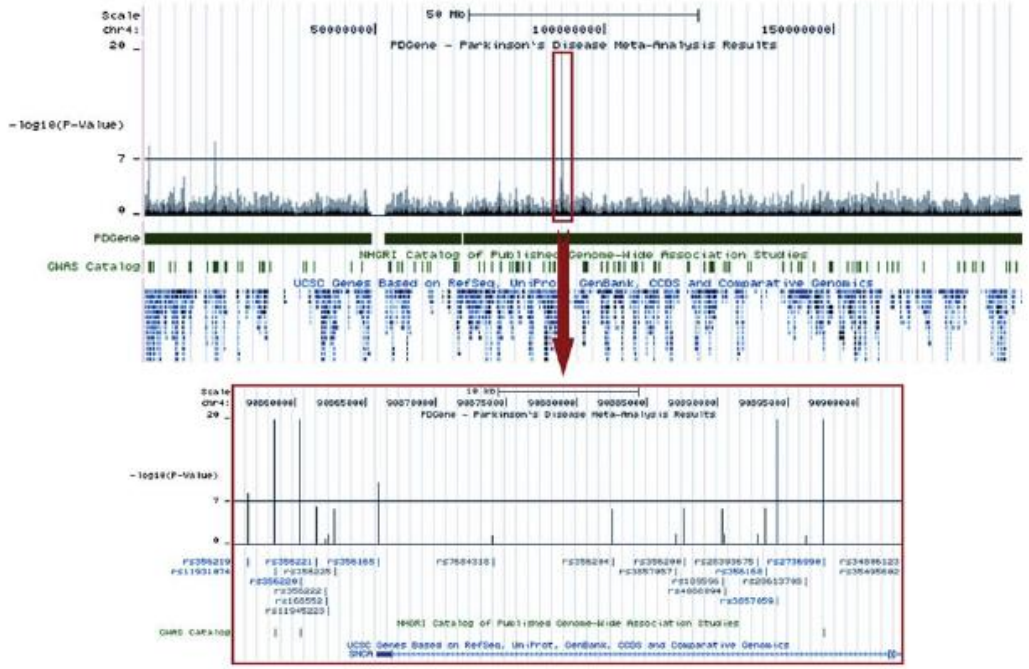


Figure 8. Example screenshot of a customized UCSC genome-browser track (PDGene database) displaying meta-analysis *p*-values. The USCS genome-browser (http://genome.ucsc.edu/cgi-bin/hgGateway) presents an up-to-date, easy-to-navigate interface of the human genome. It integrates information from a broad variety of databases (e.g. gene-mapping, expression, evolutionary conservation) at base-pair resolution. Shown is an example screenshot for a customized browser track from the PDGene database maintained by our group (http://www.pdgene.org), which summarizes meta-analysis results (-log(10)) *p*-values, based on data from both genome-wide and candidate-gene association studies) equivalently at base-pair resolution. SNPs (rs-numbers) displayed in grey are based on meta-analyses of available GWAS datasets, those displayed in blue are supplemented by candidate-gene/replication datasets. All rs-numbers cross-link back to the PDGene database to provide more details on the respective meta-analyses. Note that this feature will become available on ALSGene once initial data entry and analysis is completed.

of complex genetics: as can be seen in Figure 7, the four most recent individual datasets investigating the rs12608932 polymorphism in this gene only show very modest support in favour of a genuine risk effect (24,30,31). In particular, the study by Iida et al. (31) is interesting as this is the first to investigate this marker in non-European populations. At least in the investigated Chinese and Japanese datasets this polymorphism may have no or a much smaller role than in Europeans, although additional studies (e.g. with other markers in *UNC13A*) are needed to assess this more definitively. A similar situation of different effect size estimates across distinct ethnicities is observed for several of the current Top Results genes in our sister-database on Parkinson's disease genetics (PDGene, www.pdgene.org). Note that content for ALSGene is still under construction and that these results will probably change over time.

Future implementations

After completion of the initial curation process, all results will be ported to a customized UCSC genome-browser track displaying meta-analysis *p*-values. This will also include the results of meta-analyses on full GWAS datasets. This track will be hyperlinked to entries on ALSGene proper, where details of the sample-specific effect size estimates can be found for the most significant findings. Figure 8 shows an example screenshot of this feature for the PDGene database. In addition to displaying a browsable summary of ALSGene results on a genome-wide level, association signals can easily be put into the context of functional genetic studies, expression data, evolutionary conservation of genomic regions, etc., as annotated by the UCSC database (URL: http://genome.ucsc.edu/). This feature will also allow for an easy integration with content and results from the ALSoD database, as well as the comparison of genetic results emerging from our other database projects focusing on neurodegenerative disorders, such as Alzheimer's, Parkinson's, and frontotemporal dementia (FTDGene; in preparation).

ALSoD and ALSGene: joining forces from two different angles

As outlined above, the respective focuses of ALSoD and ALSGene lie in different aspects of genetics research in ALS. The main focus for ALSoD is the summary and curation of genotype-phenotype correlations

which is usually clearest for Mendelian ALS genes and mutational studies, while ALSGene is mainly centred on genetic association studies. Both efforts face different challenges owing to the advent of powerful high-throughput technologies now allowing a growing number of research laboratories to routinely perform whole-genome genotyping and sequencing. In ALS genetics, two of the most interesting unanswered questions to date are: 1) are the same genes causing Mendelian forms of ALS through private mutations also involved in modifying disease susceptibility of non-Mendelian forms of ALS via the action of common polymorphisms (as seems to be the case in Parkinson's disease and frontotemporal dementia); and 2) can any of the genetic associations with common polymorphisms be explained by DNA variants that are neither common nor private? While these questions can only be directly answered by performing additional laboratory experiments, the outcomes of these efforts will be tracked and fed back to the research community by ALSoD and ALSGene. Combining the emerging evidence collected in both resources via an integrated interface (e.g. using the UCSC genome-browser architecture) is one efficient, yet simple, way of achieving this goal, in addition to hyperlinking related content and results across databases. In order to accomplish these objectives, the curatorial teams of both ALSoD and ALSGene have joined forces to provide a seamless and easy-to-navigate user experience, which will hopefully allow researchers to gain new insights from apparently unrelated sets of data.

Conclusions

With the advent of increasingly more powerful technologies, what we are learning about ALS is expanding every day. This is particularly true for research in human genetics, which has provided not only ground-breaking new insights into ALS pathogenesis, but also many insubstantial results. Online research databases such as ALSoD and ALSGene serve a vital purpose in facilitating access to knowledge and generating new insights from existing disparate sets of data allowing both genetics experts and non-experts to more easily separate the wheat from the chaff in the growing mountain of raw genetic data.

Acknowledgements

ALSoD is a joint project of the World Federation of Neurology (WFN) and European Network for the Cure ALS (ENCALS). We are especially grateful for the long-standing and continued funding of this project from the ALS Association and the MND Association of Great Britain and Northern Ireland. We also thank ALS Canada, MNDA Iceland and the ALS Therapy Alliance. The research leading to these results has received funding from the European Community's Health Seventh Framework Programme FP7/2007-2013) under grant agreement n° 259867. We also thank the NIHR Biomedical Research Centre for Mental Health. We thank Aleks Radunovic, Nigel Leigh, and Ian Gowrie, who originally conceived ALSoD, and the ALSoD team, especially Peter Andersen and John Powell.

ALSGene has been funded by Prize4Life as a critical functionality of their ALS Forum research portal (www.ResearchALS.org), created to serve as a unified source of breaking ALS research news and valuable tools for the ALS research community. It has been developed in close collaboration with the Alzheimer Research Forum. We thank Gabriele Strobel, Colin Knep and Paula Noyes for the online adaptation and maintenance of ALSGene. We further thank the ALSGene database team, especially Ute Zauft, Esther Meissner, Johannes Roehr, Brit-Maren Schjeide, and Leif Schjeide.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Schottkzy J. Ueber praesenile Verbloedungen. Z Ges Neurol Psychiat. 1932;140:333–97.
2. Gowers WR. A Manual of Diseases of the Nervous System. Vol. I. Diseases of the nerves and spinal cord. 3rd edn. Philadelphia: P. Blakiston's Son & Co.; 1900.
3. Aran FA. Recherches sur une maladie non encore decrite dy systeme musculaire (atrophie musculaire progressive). Archs Gen Med. 1850;3:5–35.
4. Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature. 1993;362:59–62.
5. Hamosh A, Scott AF, Amberger J, Valle D, McKusick VA. Online Mendelian Inheritance in Man (OMIM). Hum Mutat. 2000;15:57–61.
6. Cotton RG, McKusick V, Scriver CR. The HUGO Mutation Database Initiative. Science. 1998;279:10–1.
7. Scriver CR, Nowacki PM, Lehvaslaiho H. Guidelines and recommendations for content, structure, and deployment of mutation databases: II. Journey in progress. Hum Mutat. 2000;15:13–5.
8. Claustres M, Horaitis O, Vanevski M, Cotton RG. Time for a unified system of mutation description and reporting: a review of locus-specific mutation databases. Genome Res. 2002;12:680–8.
9. George RA, Smith TD, Callaghan S, Hardman L, Pierides C, Horaitis O, et al. General mutation databases: analysis and review. J Med Genet. 2008;45:65–70.
10. Soussi T, Ishioka C, Claustres M, Beroud C. Locus-specific mutation databases: pitfalls and good practice based on the p53 experience. Nat Rev Cancer. 2006;6:83–90.
11. Radunovic A, Leigh PN. ALSODatabase: database of SOD1 (and other) gene mutations in ALS on the internet. European FALS Group and ALSOD Consortium. Amyotroph Lateral Scler Other Motor Neuron Disord. 1999;1:45–9.
12. Wroe R, Wai-Ling Butler A, Andersen PM, Powell JF, Al-Chalabi A. ALSOD: the Amyotrophic Lateral Sclerosis Online Database. Amyotroph Lateral Scler. 2008;9:249–50.
13. Yoshida M, Takahashi Y, Koike A, Fukuda Y, Goto J, Tsuji S. A mutation database for amyotrophic lateral sclerosis. Hum Mutat. 2010;31:1003–10.

14. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 2010;38: 214–20.
15. Dunckley T, Huentelman MJ, Craig DW, Pearson JV, Szelinger S, Joshipura K, et al. Whole-genome analysis of sporadic amyotrophic lateral sclerosis. N Engl J Med. 2007; 357:775–88.
16. Schymick JC, Scholz SW, Fung HC, Britton A, Arepalli S, Gibbs JR, et al. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. Lancet Neurol. 2007;6:322–8.
17. van Es MA, van Vught PW, Blauw HM, Franke L, Saris CG, Andersen PM, et al. ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. Lancet Neurol. 2007;6:869–77.
18. Cronin S, Berger S, Ding J, Schymick JC, Washecka N, Hernandez DG, et al. A genome-wide association study of sporadic ALS in a homogenous Irish population. Hum Mol Genet. 2008;17:768–74.
19. van Es MA, van Vught PW, Blauw HM, Franke L, Saris CG, van den Bosch L, et al. Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. Nat Genet. 2008;40:29–31.
20. van Es MA, Veldink JH, Saris CG, Blauw HM, van Vught PW, Birve A, et al. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. Nat Genet. 2009;41: 1083–7.
21. Landers JE, Melki J, Meininger V, Glass JD, van den Berg LH, van Es MA, et al. Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. Proc Natl Acad Sci U S A. 2009;106:9004–9.
22. Chio A, Schymick JC, Restagno G, Scholz SW, Lombardo F, Lai SL, et al. A two-stage genome-wide association study of sporadic amyotrophic lateral sclerosis. Hum Mol Genet. 2009;18:1524–32.
23. Laaksovirta H, Peuralinna T, Schymick JC, Scholz SW, Lai SL, Myllykangas L, et al. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. Lancet Neurol. 2010;9:978–85.
24. Shatunov A, Mok K, Newhouse S, Weale ME, Smith B, Vance C, et al. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. Lancet Neurol. 2010;9: 986–94.
25. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer's disease genetic association studies: the AlzGene database. Nat Genet. 2007;39:17–23.
26. Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide BM, et al. The PDGene Database. Alzheimer Research Forum. Available at: http://www.pdgene.org/.
27. Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. Int J Epidemiol. 2008;37:120–32.
28. Khoury MJ, Bertram L, Boffetta P, Butterworth AS, Chanock SJ, Dolan SM, et al. Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases. Am J Epidemiol. 2009;170:269–79.
29. Ioannidis JP. Effect of formal statistical significance on the credibility of observational associations. Am J Epidemiol. 2008;168:374–83; discussion 384–90.
30. Daoud H, Belzil V, Desjarlais A, Camu W, Dion PA, Rouleau GA. Analysis of the UNC13A gene as a risk factor for sporadic amyotrophic lateral sclerosis. Arch Neurol. 2010;67: 516–7.
31. Iida A, Takahashi A, Deng M, Zhang Y, Wang J, Atsuta N, et al. Replication analysis of SNPs on 9p21.2 and 19p13.3 with amyotrophic lateral sclerosis in East Asians. Neurobiol Aging. 2011;32:757. e13–4.

Chapter 4 Publication 3 - Credibility Analysis of Putative Disease-Causing Genes Using Bioinformatics[522]

Contribution: First Author

Journal: PLoS One

Volume: Volume 2013 Jun 5;8(6):e64899

Authors: Olubunmi Abel, John F Powell, Peter M. Andersen, Ammar Al-Chalabi

Pubmed ID: 23755159

DOI: 10.1371/journal.pone.0064899

Link: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0064899

OPEN ACCESS Freely available online

PLOS ONE

Credibility Analysis of Putative Disease-Causing Genes Using Bioinformatics

Olubunmi Abel[1], John F. Powell[2], Peter M. Andersen[3,4], Ammar Al-Chalabi[1]*

1 King's Health Partners Centre for Neurodegeneration Research, King's College London, Department of Clinical Neuroscience, London, United Kingdom, **2** Department of Neuroscience, King's College London, London, United Kingdom, **3** Institute of Pharmacology and Clinical Neuroscience, Section for Neurology, Umeå University, Umeå, Sweden, **4** Department of Neurology, University of Ulm, Ulm, Germany

Abstract

Background: Genetic studies are challenging in many complex diseases, particularly those with limited diagnostic certainty, low prevalence or of old age. The result is that genes may be reported as disease-causing with varying levels of evidence, and in some cases, the data may be so limited as to be indistinguishable from chance findings. When there are large numbers of such genes, an objective method for ranking the evidence is useful. Using the neurodegenerative and complex disease amyotrophic lateral sclerosis (ALS) as a model, and the disease-specific database ALSoD, the objective is to develop a method using publicly available data to generate a credibility score for putative disease-causing genes.

Methods: Genes with at least one publication suggesting involvement in adult onset familial ALS were collated following an exhaustive literature search. SQL was used to generate a score by extracting information from the publications and combined with a pathogenicity analysis using bioinformatics tools. The resulting score allowed us to rank genes in order of credibility. To validate the method, we compared the objective ranking with a rank generated by ALS genetics experts. Spearman's Rho was used to compare rankings generated by the different methods.

Results: The automated method ranked ALS genes in the following order: *SOD1*, *TARDBP*, *FUS*, *ANG*, *SPG11*, *NEFH*, *OPTN*, *ALS2*, *SETX*, *FIG4*, *VAPB*, *DCTN1*, *TAF15*, *VCP*, *DAO*. This compared very well to the ranking of ALS genetics experts, with Spearman's Rho of 0.69 ($P = 0.009$).

Conclusion: We have presented an automated method for scoring the level of evidence for a gene being disease-causing. In developing the method we have used the model disease ALS, but it could equally be applied to any disease in which there is genotypic uncertainty.

Citation: Abel O, Powell JF, Andersen PM, Al-Chalabi A (2013) Credibility Analysis of Putative Disease-Causing Genes Using Bioinformatics. PLoS ONE 8(6): e64899. doi:10.1371/journal.pone.0064899

Editor: Bart Dermaut, Pasteur Institute of Lille, France

Received January 24, 2013; **Accepted** April 19, 2013; **Published** June 5, 2013

Funding: The authors are especially grateful for the long-standing and continued funding of this project from the ALS Association and the MND Association of Great Britain and Northern Ireland. They also thank ALS Canada, MNDA Iceland and the ALS Therapy Alliance for support. The research leading to these results has received funding from the European Community 's Health Seventh Framework Programme FP7/2007–2013 under grant agreement number 259867. AA-C receives salary support from the National Institute for Health Research (NIHR) Dementia Biomedical Research Unit at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. Aleks Radunovic, Nigel Leigh, and Ian Gowrie originally conceived ALSoD. ALSoD is a joint project of the World Federation of Neurology (WFN) and European Network for the Cure ALS (ENCALS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: AA-C is a consultant for Biogen Idec and Cytokinetics. There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: al-chalabi@kcl.ac.uk

Introduction

Genetic studies are challenging in many complex diseases, particularly those with limited diagnostic certainty, low incidence and prevalence, or those of old age. Association studies suffer a reduction in power when there is phenotypic heterogeneity resulting from difficulty with diagnosis, and linkage studies are limited because the older generations are not available and the younger generations have not yet reached the age of risk. The result is that genes are reported as causative with varying levels of evidence and it can be difficult for those not in the field to assess how credible any genetic evidence is.

One such condition is amyotrophic lateral sclerosis (ALS). This is an adult onset neurodegenerative syndrome of upper and lower motor neuron degeneration, with a mean age of onset of 56 in diagnosed familial cases (FALS) and 60 to 70 years in apparently sporadic cases, and an average survival of 3 to 5 years from symptom onset [1] [2]. Illustrating the complexity and difficulty in performing genetic research on ALS, the reported frequency of familial ALS varies from 0.8% [3] to 17–18% [4] although all studies agree that most cases are apparently sporadic [5]. There is, however, a genetic basis both to familial and apparently sporadic ALS [6,7,8]. All genes reported mutated in familial ALS have also been found mutated in sporadic ALS. Because of the late age of onset and poor prognosis, suitable families are difficult to collect for linkage, and large populations are difficult to collect for association.

The first gene identified for familial ALS was *SOD1* [9] [10]. Through linkage and association studies of SNPs, microsatellites and copy number variants, as well as through direct sequencing of candidate genes and whole exome sequencing using high throughput methods, over 100 genes have now been implicated in the cause of ALS [11]. The level of supporting evidence for each gene or gene variant varies from small to overwhelming, and is in some cases contradictory. Furthermore, the increasing cooperation between ALS researchers internationally, and the understanding that large datasets are needed, coupled with advances in technology, mean that the rate of detection of putative new ALS genes is rapid and increasing. This leads to two immediate problems: first, it is difficult to keep up with what is an "accepted" ALS gene, and second, there is no simple, objective way to define the list of ALS-causing genes. As a result, researchers may find themselves unable to agree on whether any one gene is an ALS gene or not. The situation is further compounded by the loose definition of ALS, which for genetic purposes has a far wider phenotypic definition than most ALS researchers would accept in a clinical setting [12]. For example, ALS2 includes an infantile, slowly progressive upper motor neuron syndrome that is most similar to hereditary spastic paraparesis, rather than an adult onset mixed upper and lower motor neuron syndrome with a poor prognosis for survival. Similarly, ALS with frontotemporal dementia is regarded as a slightly different entity from ALS even though frontotemporal dementia and ALS are in at least some cases a continuum of disease, and in many cases ALS genes and frontotemporal dementia genes are the same as genes for ALS with frontotemporal dementia.

One solution to this problem is to design some method for objectively scoring the level of evidence supporting a gene or gene variant as disease causing. This would have the advantage that the phenotype could be defined by the user, allowing a loose definition or more stringent definition as required.

The ALSoD database stores data on putative ALS genes using information derived from publications and directly input by researchers. We have therefore explored the possibility of using these data to generate a credibility score for ALS genes with the aim of producing a system that can be generalized to other similar conditions.

Methods

PRISMA revision [13] with respect to development and reporting of results were taken into consideration. (Checklist S1).

Data Collection

Genes with at least one publication suggesting involvement in adult onset familial ALS were studied [14]. We excluded genes with limited clinical data, absent mutational data or unreplicated results. Publicly listed variants for the included genes derived from ALSGene, Uniprot, ALS Mutation and HGMD databases were merged with variant lists in ALSoD, and filtered for duplicates (Figure 1).

Pathogenicity Analysis Using Bioinformatic Tools

PANTHER (Protein Analysis Through Evolutionary Relationships) [15], SIFT (Sorting Intolerant From Tolerant) [16] and POLYPHEN (Polymorphism Phenotyping) [17] programs were used to analyse variants for possible pathogenicity. These tools generated a set of scores for the variants analysed, which for PANTHER are given as a subSPEC (substitution position-specific evolutionary conservation) score and for POLYPHEN given as score differences for PSIC (position-specific independent counts). In PANTHER, all possible mutations for each gene were generated using perl scripts and run on the web service in batches. SubPSEC scores $<=-5.0$ were defined as damaging and subPSEC scores >-5.0 defined as not damaging. In SIFT, all possible unique codons in each gene were generated using perl script with scores $<=0.05$ defined as damaging and scores >0.05 defined as not damaging. In POLYPHEN, all mutations available in a gene on ALSoD were run through the web service one after the other and PSIC score differences $>=1.5$ defined as damaging and PSIC score differences <1.5 as not damaging.

Data Extraction from Publications

We conducted a systematic review of all publications related to ALS genetics with an exhaustive combination of search queries on the 15 genes mentioned above. (Flow diagram S1 and Protocol S1).

In the PubMed database, we used title keywords consisting of the gene name, "mutation" and "ALS" or "Motor Neuron Disease", or gene name and "novel" to identify key publications and then used the related citations function to generate a list of publications for data extraction. For example, (SOD1[Title] OR (superoxide dismutase[Title]) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) yielding 181 results. These results were further filtered by choosing "Humans" as Species and sorted by "Recently Added" thereby displaying 160 unique publications. From the list displayed, we also searched the "Related citations" link on the first publication [10] of the selected gene SOD1 yielding 204 results.

We used Google Scholar (http://scholar.google.co.uk/) to identify publications for import into the ALSoD database, starting with basic search queries to generate a large number of publications. For example, "SOD1" gave about 28,600 results but "SOD1 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" gave 2050 results. We went through the first 20 pages containing 20 publications on each page and already sorted by relevance. Publications with animal models or associated with other diseases were excluded from the long list. A manual comparison with already discovered publications from pubmed was conducted and these were excluded from the list.

Manually curated data extracted from all publications included family history, El Escorial category [18,19] mutations per gene, number of cases and controls used in the studies, mutations in the same codon, number of patients with family history (FALS), number of patients without family history, mutations replicated in other studies, number of countries replicating the mutation and for linkage studies, LOD scores. Several genes implicated in ALS are also implicated in other diseases, including frontotemporal dementia, spinocerebellar ataxia and parkinsonism. To avoid the problem of non-ALS patients being included in the database, we restricted data curation to publications specifying ALS.

Automated Gene Ranking

Eleven queries stored as procedures were performed on data collated. These were: 1. The total number of affected patients with El Escorial defined ALS having a mutation in each gene [14,15]. 2. The total number of ALS affected patients used in each study. This measure was used to account for sampling variance and power [20]. 3. The total number of healthy individuals with a mutation reported in each study. 4. The total number of healthy individuals used in each study. 5. The total number of mutations sharing the same codon. 6. The total number of variants detected in ALS patients for each gene. 7. The total number of mutations with positive pathogenic predictions from the use of the three

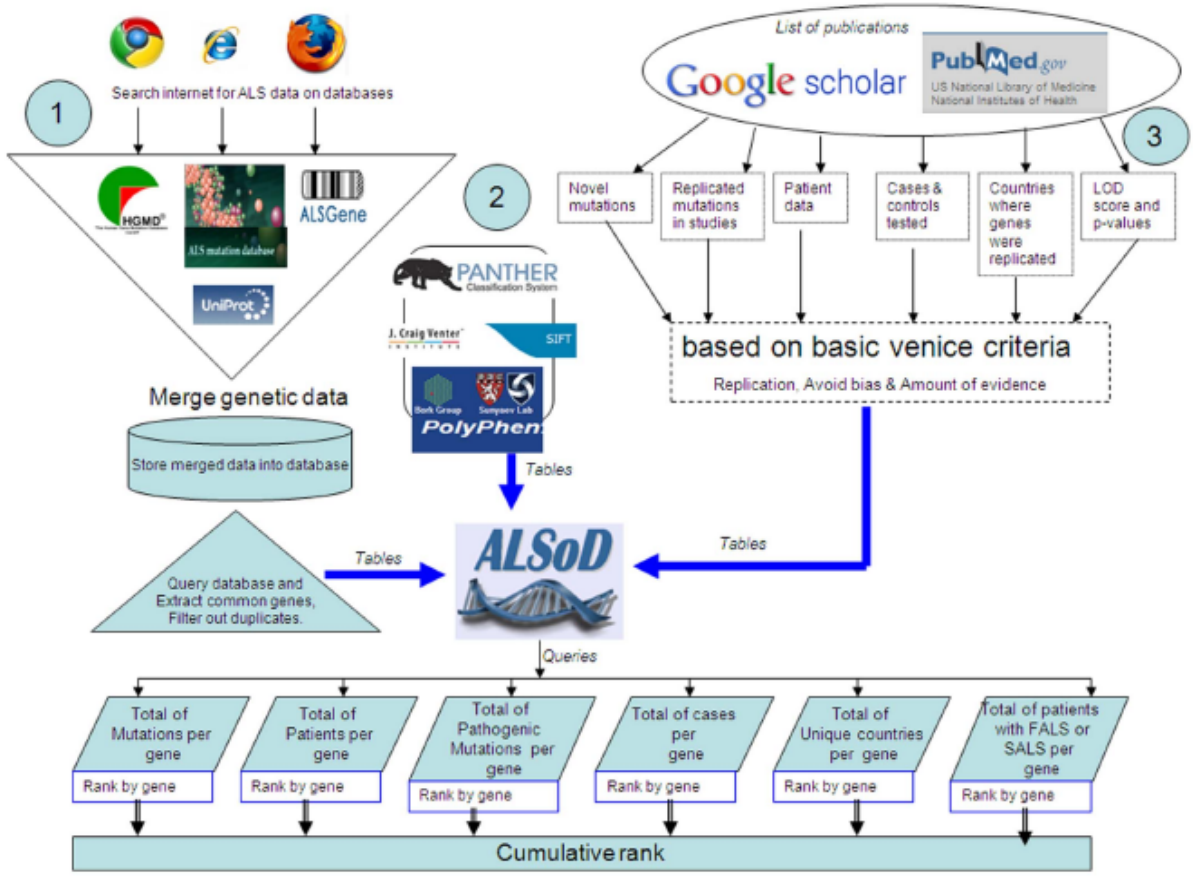


Figure 1. Overview of credibility analysis method.
doi:10.1371/journal.pone.0064899.g001

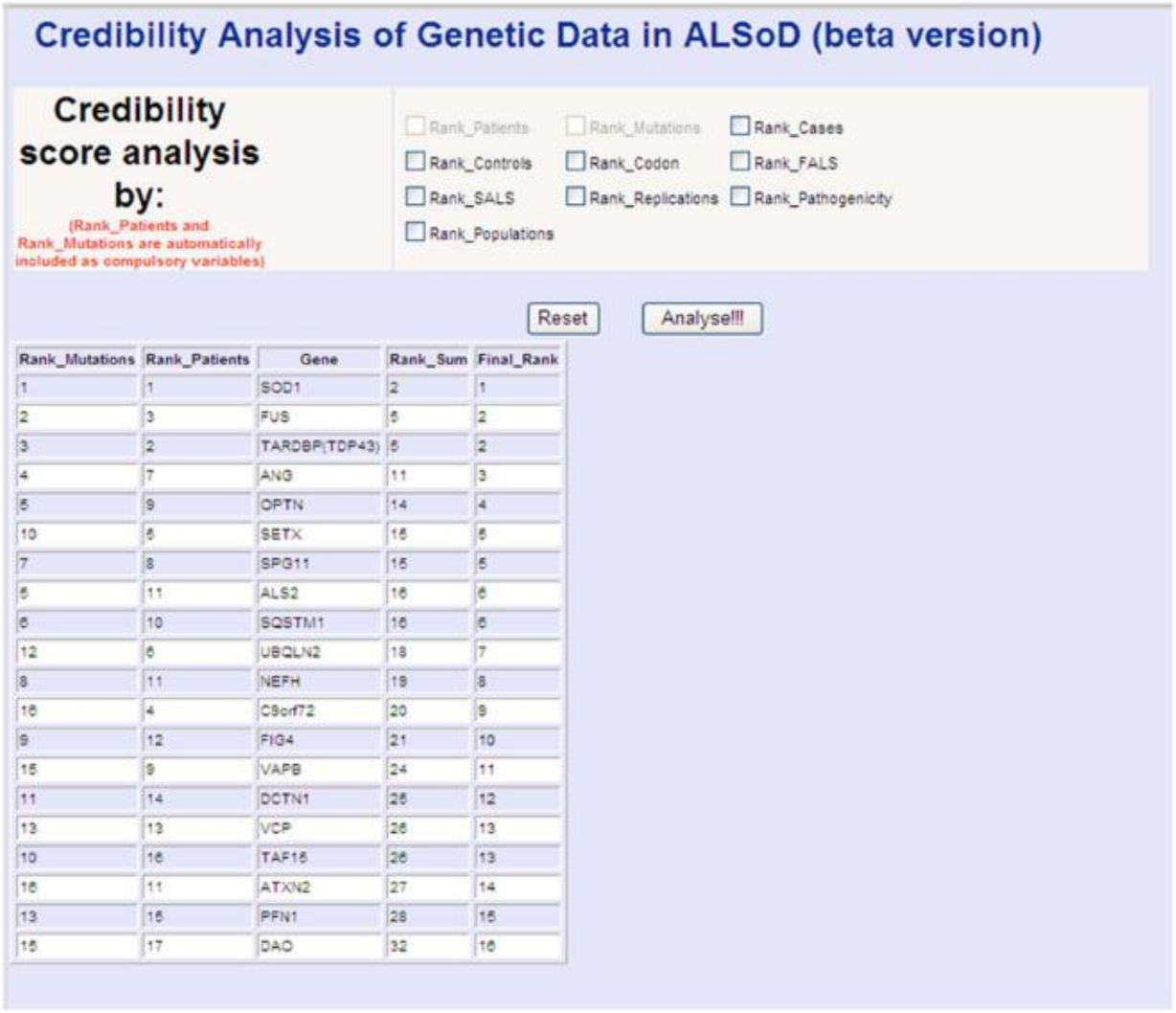
bioinformatics tools described above. 8. The number of patients with a family history defined as at least one other affected member of the family. 9. The number of patients without a family history of ALS. 10. The number of times a particular variation was replicated across different studies. 11. The number of unique populations where affected patients originated.

For each procedure above, a query was generated using Structured Query Language (SQL) on Microsoft SQL Server 2008 and displayed on the ASP.NET platform webpage, ranking the gene. The predicted pathogenicity score for each tool was scored 1 for predicted pathogenic and 0 for predicted not pathogenic and then summed to generate a final score for ranking (http://alsod.iop.kcl.ac.uk/Statistics/pathogenicity.aspx). The rank score for each query was summed to generate an overall rank for the gene under study. For example, from Figure 2, the last row for the *DAO* gene gives the column score 15 for Rank_Mutations, 14 for Rank_Patients and 9 for Rank_Pathogenicity. This produces a total of 38 (that is 15+14+9) in the Rank_Sum column. The generated Rank_Sum for all the genes are arranged in ascending order placing *DAO* 12th by final rank. On the other hand, *FUS* is placed 3rd by final rank as the corresponding scores are 3+3+2 = 8.

There are two possible ways of ranking results in SQL. The default method allocates rank based on the true position, such that if two genes are given equal first position for example, the next gene is in third position, not second. The dense rank method allocates the next gene as second so that there are no gaps in the rank numbering. We used the dense ranking system.

Validation of the Method

The purpose of the credibility score tool is to generate a list of genes in order of the weight of evidence supporting involvement in ALS. Such a list should correlate closely with one generated by ALS genetics experts, since such experts should have a good working knowledge of the available evidence. We therefore conducted a survey of ALS genetic experts, defined as being individuals who had published as first or senior author on ALS genetics. Experts were surveyed using the freely available online questionnaire tool, Surveymonkey on http://www.surveymonkey.com/s/WRDW5WT (Figure 3). The survey link showed the genes randomly ordered differently every time the link was clicked to prevent bias in the responses that might occur based on ordering. We also embedded the questionnaire as a submenu on the feedback menu of the ALSoD website. Experts were randomly assigned to one of two groups, one in which the same rank could be assigned to several genes, and one in which responders were forced to rank each gene in order. The first group mimics the final score of the automated method closely, while the second group mimics the detail of the automated ranking method closely, since the automated method is forced to rank each query uniquely but



Rank_Mutations	Rank_Patients	Gene	Rank_Sum	Final_Rank
1	1	SOD1	2	1
2	3	FUS	5	2
3	2	TARDBP(TDP43)	5	2
4	7	ANG	11	3
5	9	OPTN	14	4
10	6	SETX	16	5
7	8	SPG11	15	5
5	11	ALS2	16	6
6	10	SQSTM1	16	6
12	6	UBQLN2	18	7
8	11	NEFH	19	8
16	4	C9orf72	20	9
9	12	FIG4	21	10
15	9	VAPB	24	11
11	14	DCTN1	25	12
13	13	VCP	26	13
10	16	TAF15	26	13
16	11	ATXN2	27	14
13	15	PFN1	28	15
15	17	DAO	32	16

Figure 2. Credibility Analysis webpage.
doi:10.1371/journal.pone.0064899.g002

the combined ranking could result in the same value for different genes.

User Interface

The Credibility Analysis page at (http://alsod.iop.kcl.ac.uk/Statistics/credibility.aspx) allows criteria to be selected by users in the form of checkboxes. Clicking the 'Analyse' button then displays the ranked result. A detailed summary of ranked credibility data are also displayed for further reference by users giving the outcome of each procedure based query. Any combination of queries can be included in generating the score except Number of patients and Number of mutations found in each gene which are mandatory selections.

Statistical Methods

Spearman's Rho [21,22,23] was used to compare rankings generated by the automated method and the ALS genetics experts.

Results

For the pathogenicity prediction, using a threshold score >1 (that is, where the combination score is 2 or 3) to define pathogenicity, just 110 mutations out of 425 were identified as pathogenic, with particularly poor predictions for *FUS* and *TARDBP* when compared with biological evidence of pathogenicity. Using a threshold score of >0 (that is, where the combination score is 1 or 2 or 3) to define pathogenicity brought the number of pathogenic mutations to 198, suggesting that about 50% of recorded FALS mutations are pathogenic based on bioinformatics predictions.

There were 14 genes that fulfilled the inclusion criteria for generation of a credibility score at the time of the survey, and had sufficient data manually curated from publications as explained in the data extraction process above. These were *ALS2*, *FUS*, *DAO*, *VCP*, *VAPB*, *ANG*, *DCTN1*, *FIG4*, *SETX*, *SOD1*, *TARDBP*, *SPG11*, *NEFH*, and *OPTN*.

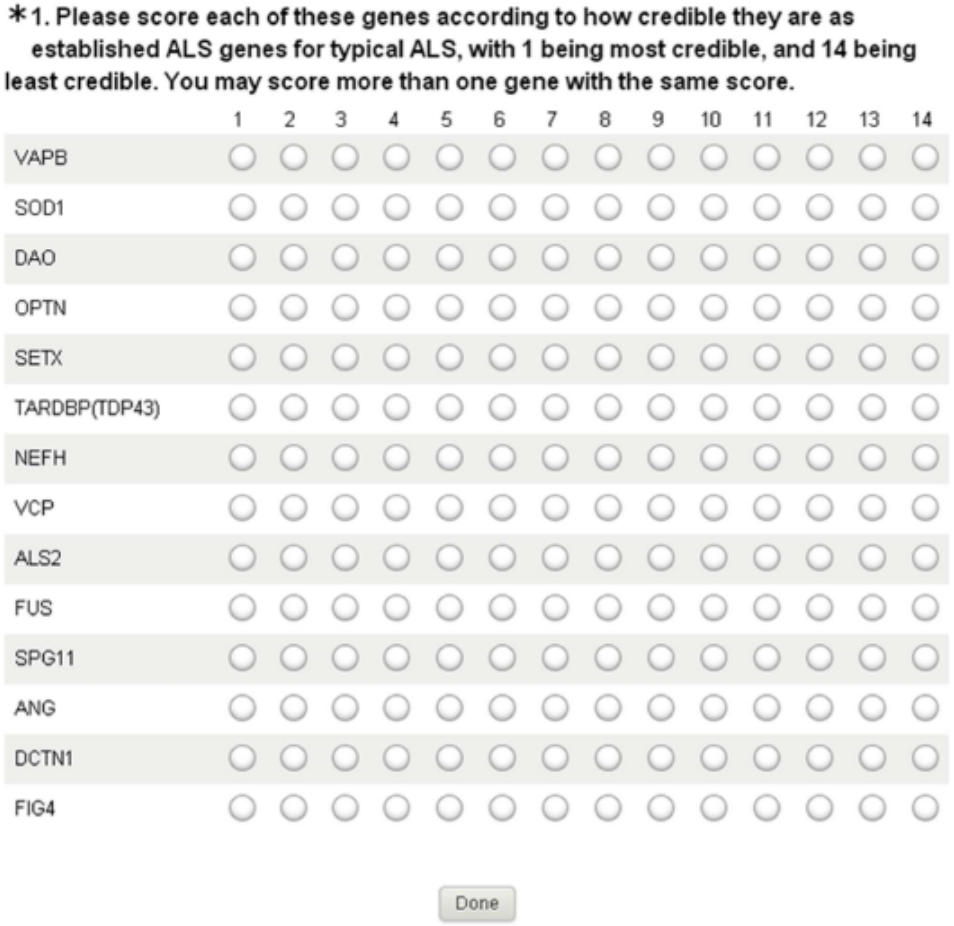


Figure 3. Surveymonkey survey tool for ranking 14 genes.
doi:10.1371/journal.pone.0064899.g003

Using the full set of 11 procedures, the automated method ranked these as ALS-causing genes in the following order: *SOD1*, *TARDBP*, *FUS*, *ANG*, *SPG11*, *NEFH*, *OPTN*, *ALS2*, *SETX*, *FIG4*, *VAPB*, *DCTN1*, *TAF15*, *VCP*, *DAO*.

Subsets of the 11 procedures may be defined by the user if needed. This allows flexibility in which evidence is regarded as useful. For example in Figure 3, using the number of mutations reported in a single gene and the number predicted as pathogenic as test criteria ranks the genes in the following order: *SOD1*, *TARDBP*, *FUS*, *ANG*, *OPTN*, *SETX*, *ALS2*, *SPG11*, *FIG4*, *DCTN1*, *VAPB*, *VCP*, *DAO*. The output shows that the first six genes, *SOD1*, *TARDBP*, *FUS*, *ANG*, *OPTN* and *SETX*, have a total of 121, 17, 19, 12, 5 and 4 pathogenic mutations respectively and, for example, the I113T, D90A and A4V pathogenic mutations of the *SOD1* gene were replicated in 17, 14 and 12 studies. It also shows there are 6 different mutations in codon 93 of *SOD1* and 5 different mutations in codon 521 of *FUS*. Other displayed information includes the number of countries in which gene mutations have been reported. For example, *SOD1* mutation has been reported in 34 countries with representation from every continent of the world, while *TARDBP*, *ALS2*, *ANG*, *FUS*, *SETX* and *NEFH* have been reported in 13, 9, 7, 7, 6 and 5 unique countries respectively. Genes like *FIG4*, *DPP6*, *DCTN1*, *UBQLN2*, *TAF15* which were recorded in only 1 country each have the lowest ranks.

8/25 ALS genetics experts selected based on having published at least one paper on ALS genetics responded. Comparison of the full automated method with the ALS genetics experts' rankings gave a Spearman's Rho of 0.69 ($P = 0.009$) for the forced expert rankings, and 0.57 ($P = 0.042$) for the unforced rankings, indicating a good correlation between the methods.

Discussion

We have presented an automated method for using published information to score the level of evidence supporting a causative relationship between gene mutation and a disease. The information on which the credibility analysis is based is collected routinely by locus-specific databases and the method can therefore be generalized to other diseases. The method used has been applied to amyotrophic lateral sclerosis but could equally be applied to any disease in which there is phenotypic and genotypic heterogeneity.

A strength of this method is that multiple lines of evidence are used to generate an objective opinion as to the credibility of a gene as a disease gene, and while publication bias will affect the score, this is minimized by several factors. First, in this study unpublished

data are used since the database includes directly input information from researchers who have not published. Second, a major part of the score is generated using theoretical models of pathogenicity. Third, once published, any information remains useable, and not prone to the vagaries of scientific fashion, or the bias of individual opinion leaders. The effects of these components on the score can be seen by comparing the automated ranking and the ranking generated by both groups of ALS genetics experts. In general the rankings were in agreement. For example, with one exception, the top five genes were the same for all three methods. For some genes there were strikingly different ranks. ANG was ranked 9 of 13 by the experts who could give equal ranks, but in the top five for the other two methods. The biggest discrepancies were otherwise for ALS2, NEFH, and VAPB, each of which was ranked in the bottom two for one of the methods and in the middle for the other two methods.

Similar approaches have been used in association studies. In previous work, three criteria used to determine how credible a disease gene might be were the amount of evidence, manifest as number of studies and population size studied, replicability of a result, and protection from bias by good study design [24]. We have tried to follow similar principles in generating this credibility score.

A weakness of this method is that it relies on an agreed set of criteria for analysis to generate the score, but there is no way to decide objectively whether the criteria are reasonable or what their relative weights should be. For example, we have not included pathogenicity demonstrated in animal models in the score but others might regard this as a vital component. Although we have tried to build in flexibility so that researchers can include or exclude certain criteria, unless the available criteria are exhaustive there will always be the possibility that the method is incomplete. Similarly, because the criteria can be user-selected, there can be no truly universal measure of credibility using this system.

Since this tool was developed, pathological expansion in the *C9orf72* gene has been identified as a cause of ALS and frontotemporal dementia [25,26]. At the time of our survey of experts this was not the case and it has therefore been excluded from the analysis presented.

A major advantage of this tool is the automation which changes the rank of a gene depending on the evidence provided on the database. This system could be applied to other complex diseases where multiple genes are responsible for a phenotype.

Supporting Information

Checklist S1
(DOC)

Flow Diagram S1
(DOC)

Protocol S1
(DOC)

Author Contributions

Conceived and designed the experiments: OA JFP PMA AA-C. Wrote the paper: OA JFP PMA AA-C. Advised on criteria used and literature review: JFP PMA AA-C. Contributed genetic data: PMA AA-C. Survey and Statistical Analysis of data: OA AA-C.

References

1. Charcot J, Joffroy A (1869) Deux cas d'atrophie musculaire progressive. Arch Physiol Norm Pathol 2: 744–760.
2. Cleveland DW, Rothstein JD (2001) From Charcot to Lou Gehrig: deciphering selective motor neuron death in ALS. Nature Reviews Neuroscience 2: 806–819.
3. Fong GCY, Kwok KHH, Song Y, Cheng T, Ho PWL, et al. (2006) Clinical phenotypes of a large Chinese multigenerational kindred with autosomal dominant familial ALS due to Ile149Thr SOD1 gene mutation. Amyotrophic Lateral Sclerosis 7: 142–149.
4. Eisen A, Mezei MM, Stewart HG, Fabros M, Gibson G, et al. (2008) SOD1 gene mutations in ALS patients from British Columbia, Canada: clinical features, neurophysiology and ethical issues in management. Amyotrophic Lateral Sclerosis 9: 108–119.
5. Byrne S, Walsh C, Lynch C, Bede P, Elamin M, et al. (2011) Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis. Journal of Neurology, Neurosurgery and Psychiatry 82: 623.
6. Al-Chalabi A, Lewis CM (2011) Modelling the effects of penetrance and family size on rates of sporadic and familial disease. Human Heredity 71: 281–288.
7. Hanby MF, Scott KM, Scotton W, Wijesekera L, Mole T, et al. (2011) The risk to relatives of patients with sporadic amyotrophic lateral sclerosis. Brain.
8. Al-Chalabi A, Leigh PN (2005) Trouble on the pitch: are professional football players at increased risk of developing amyotrophic lateral sclerosis? Brain 128: 451.
9. Siddique T, Figlewigz DA, Pericak-Vance MA, Haines JL, Rouleau G, et al. (1991) Linkage of a gene causing familial amyotrophic lateral sclerosis to chromosome 21 and evidence of genetic-locus heterogeneity. New England Journal of Medicine 324: 1381–1384.
10. Rosen D, Siddique T, Patterson D, Figlewicz D, Sapp P, et al. (1993) Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis.
11. Lill CM, Abel O, Bertram L, Al-Chalabi A (2011) Keeping up with genetic discoveries in amyotrophic lateral sclerosis: The ALSoD and ALSGene databases. Amyotrophic Lateral Sclerosis 12: 238–249.
12. Hamosh A, Scott A, Amberger J, Valle D, McKusick V (2000) Online Mendelian Inheritance in Man (OMIM) Hum. Mutat 15: 57–61.
13. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine 6: e1000097.
14. Andersen PM, Al-Chalabi A (2011) Clinical genetics of amyotrophic lateral sclerosis: what do we really know? Nature Reviews Neurology.
15. Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, et al. (2003) PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. Nucleic Acids Research 31: 334.
16. Ng PC, Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Research 31: 3812.
17. Sunyaev S, Ramensky V, Koch I, Lathe W III, Kondrashov AS, et al. (2001) Prediction of deleterious human alleles. Human Molecular Genetics 10: 591.
18. Brooks BR (1994) El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. Journal of the Neurological Sciences 124: 96.
19. Brooks BR, Miller RG, Swash M, Munsat TL (2000) El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotrophic lateral sclerosis and other motor neuron disorders: official publication of the World Federation of Neurology, Research Group on Motor Neuron Diseases 1: 293.
20. Agency for Toxic Substances and Disease Registry (2011) National Amyotrophic Lateral Sclerosis (ALS) Registry.
21. David F, Mallows C (1961) The variance of Spearman's rho in normal samples. Biometrika 48: 19–28.
22. Ramsey PH (1989) Critical values for Spearman's rank order correlation. Journal of Educational and Behavioral Statistics 14: 245–253.
23. Yue S, Pilon P, Cavadias G (2002) Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. Journal of Hydrology 259: 254–271.
24. Ioannidis J, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, et al. (2008) Assessment of cumulative evidence on genetic associations: interim guidelines. International Journal of Epidemiology 37: 120.
25. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, et al. (2011) Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. Neuron.
26. Renton AE, Majounie E, Waite A, SimÃ3n-SÃ¡nchez J, Rollinson S, et al. (2011) A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. Neuron.

Chapter 5 Publication 4 - Development of a Smartphone App for a Genetics Website: The Amyotrophic Lateral Sclerosis Online Genetics Data [523]

Contribution: First Author

Journal: JMIR mHealth and uHealth

Volume: Sep2013, Vol. 15 Issue 9, p1-1. 1p. 2

Authors: Olubunmi Abel, John F Powell, Peter M. Andersen, Ammar Al-Chalabi

Pubmed ID: Not yet

DOI: 10.2196/mhealth.2706

Link: http://mhealth.jmir.org/2013/2/e18/

Original Paper

Development of a Smartphone App for a Genetics Website: The Amyotrophic Lateral Sclerosis Online Genetics Database (ALSoD)

Olubunmi Abel[1]; Aleksey Shatunov[1]; Ashley R Jones[1]; Peter M Andersen[2]; John F Powell[3]; Ammar Al-Chalabi[1]

[1]King's College London, Institute of Psychiatry, Department of Clinical Neuroscience, London, United Kingdom

[2]Department of Pharmacology and Clinical Neuroscience, Section for Neurology, Umeå University, SE-901 85 Umeå, Sweden

[3]King's College London, Department of Neuroscience, London, United Kingdom

Corresponding Author:
Ammar Al-Chalabi
King's College London
Institute of Psychiatry
Department of Clinical Neuroscience
De Crespigny Road, IoP Main building, 4th Floor
London, SE5 8AF
United Kingdom
Phone: 44 20 7848 5187
Fax: 44 20 7848 5190
Email: ammar.al-chalabi@kcl.ac.uk

Abstract

Background: The ALS Online Genetics Database (ALSoD) website holds mutation, geographical, and phenotype data on genes implicated in amyotrophic lateral sclerosis (ALS) and links to bioinformatics resources, publications, and tools for analysis. On average, there are 300 unique visits per day, suggesting a high demand from the research community. To enable wider access, we developed a mobile-friendly version of the website and a smartphone app.

Objective: We sought to compare data traffic before and after implementation of a mobile version of the website to assess utility.

Methods: We identified the most frequently viewed pages using Google Analytics and our in-house analytic monitoring. For these, we optimized the content layout of the screen, reduced image sizes, and summarized available information. We used the Microsoft .NET framework mobile detection property (HttpRequest.IsMobileDevice in the Request.Browser object in conjunction with HttpRequest.UserAgent), which returns a true value if the browser is a recognized mobile device. For app development, we used the Eclipse integrated development environment with Android plug-ins. We wrapped the mobile website version with the WebView object in Android. Simulators were downloaded to test and debug the applications.

Results: The website automatically detects access from a mobile phone and redirects pages to fit the smaller screen. Because the amount of data stored on ALSoD is very large, the available information for display using smartphone access is deliberately restricted to improve usability. Visits to the website increased from 2231 to 2820, yielding a 26% increase from the pre-mobile to post-mobile period and an increase from 103 to 340 visits (230%) using mobile devices (including tablets). The smartphone app is currently available on BlackBerry and Android devices and will be available shortly on iOS as well.

Conclusions: Further development of the ALSoD website has allowed access through smartphones and tablets, either through the website or directly through a mobile app, making genetic data stored on the database readily accessible to researchers and patients across multiple devices.

(JMIR Mhealth Uhealth 2013;1(2):e18) doi:10.2196/mhealth.2706

KEYWORDS

ALSoD; amyotrophic lateral sclerosis; frontotemporal dementia; Web-bases; database; genetics; bioinformatics; mobile website; app

Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of motor neurons resulting in progressive weakness of voluntary muscles. Death usually follows 2-5 years after the first symptoms appear, due to respiratory failure [1]. The causes of ALS are largely unknown, but a genetic component is present even in those without a family history of the disease. The gene variants contributing to risk have been identified in about 15% of the affected population, and the proportion is increasing rapidly. The ALS Online Genetics Database (ALSoD) [2] is a freely available database funded by major ALS charities (the ALS Association, the Motor Neurone Disease Association, ALS Canada, and MNDA Iceland) and sponsored by the European Network for the Cure of ALS and the World Federation of Neurology. It is aimed at researchers and clinicians and is designed for collation and bioinformatics analysis of ALS gene and phenotypic information. It typically receives about 300 unique visits a day [3,4].

Like most websites, ALSoD was initially built to display information to users of desktops and laptops, and the pages are configured to suit the height and width of those screens. With changes in browsing habits, website access is now often through a small portable device like a smartphone or tablet computer. It is therefore essential to ensure data will display correctly on a small device [5,6]. Mobile device data traffic has overtaken desktop traffic in the last decade, and data traffic on mobile devices for browsing alone has risen more than four times in 2008 [5]. According to NetMarketShare, the introduction of the Apple mobile device operating system, iOS, for the Apple iPad and iPhone for mobile browsing between March and October 2010 doubled Internet traffic [7], leading to a projection that, by 2014, mobile Internet usage will overtake desktop Internet usage [8]. Although the target community for ALSoD is mainly university or hospital-based users where desktop and laptop computers are common, such users are increasingly likely to use a portable device for use in clinic settings, conferences, or the laboratory, where fast access away from an office may be needed. Thus, it is essential that the ALSoD website is accessible not just from a desktop or laptop computer, but also from portable devices.

Mobile phone displays have greatly improved from the early monochrome screens for sending SMS messages to colorful graphical touch screens for mobile browsing [9]. Issues of low bandwidth and low resolution screens have been resolved, and smartphones and tablets should be regarded as "mobile computers" [10-13]. It therefore makes sense to write webpages specifically for portable devices. Mobile webpage content is similar to desktop webpage content and uses HTML connected and accessible over the Internet, even though mobile websites are typically accessed through Wi-Fi, 3G, or 4G networks [14]. Furthermore, portable devices are able to use applications (apps) specific to a website for access, rather than a generic browser, with the advantage of offline access to some information. Apps are generally platform-specific and downloaded from company portals, for example, BlackBerry App World, Apple App Store, and Google Play [14]. Thus, as well as writing ALSoD webpages specific for mobile devices, we also aimed to design a platform-specific app to enable some offline content and improve the user experience.

Methods

Optimization of Webpages

Our first focus was on the development of a mobile Web-based platform because we wanted the content to work across all mobile platforms [6]. Identification of the most viewed information was carried out using in-house analytic data coupled with the Google Analytics service configured for the ALSoD website. The Google Analytics tool was configured in August 2012. We based our data analysis on the 3 subsequent months (August, September, and October). These represent the period from when the Google Analytics tool was implemented to the point where we started the development of the mobile website and will be referred to as the pre-mobile website period. The 3-month period from November 2012 to January 2013 with the mobile website fully developed and the app implemented will be referred to as the post-mobile website period. In the pre-mobile website period, page views (the total number of pages viewed) were analyzed to discover the commonly visited pages on the website, including repeated views of a single page.

Design Heuristics

Designing a mobile website that works on several platforms does not mean shrinking a complete webpage into a mini-sized webpage; the aim is to eliminate very small type, scrolling left or right, and typing, and to achieve the outcome required with a single click [6,15]. We therefore created a separate style sheet, retained some original images, reduced the size of images by a fixed percentage, configured the content layout of the screen to wrap text, and summarized the information on the desktop version to fit a smaller screen.

Mobile Device Detection

To detect if the request comes from a mobile device, the .NET Framework provides the "isMobileDevice" property, which returns a true value if the browser is a recognized mobile device. This does not always work because some mobile device browsers disguise themselves as desktop browsers [16]. We therefore use "UserAgent" strings sent by the mobile device browser to the server in conjunction with the "isMobileDevice", as described in Multimedia Appendix 1 and with an overview in Figure 1.

Requesting Responses

Text messages and BlackBerry Messenger messages were sent to a selection of individuals known to the authors, asking them to view the ALSoD Web address on their phones and tablets. After the app was developed, additional users were asked to download the app via Google Play. Since ALSoD has a Facebook [17] account, a Facebook "Recommend" Button was embedded on the mobile master page. All users were asked to click on the Facebook Recommend button so as to have an estimate of the number of users who were satisfied with the outcome of the display on their phone, as seen in Figure 2.

Figure 1. Overview of mobile website development.

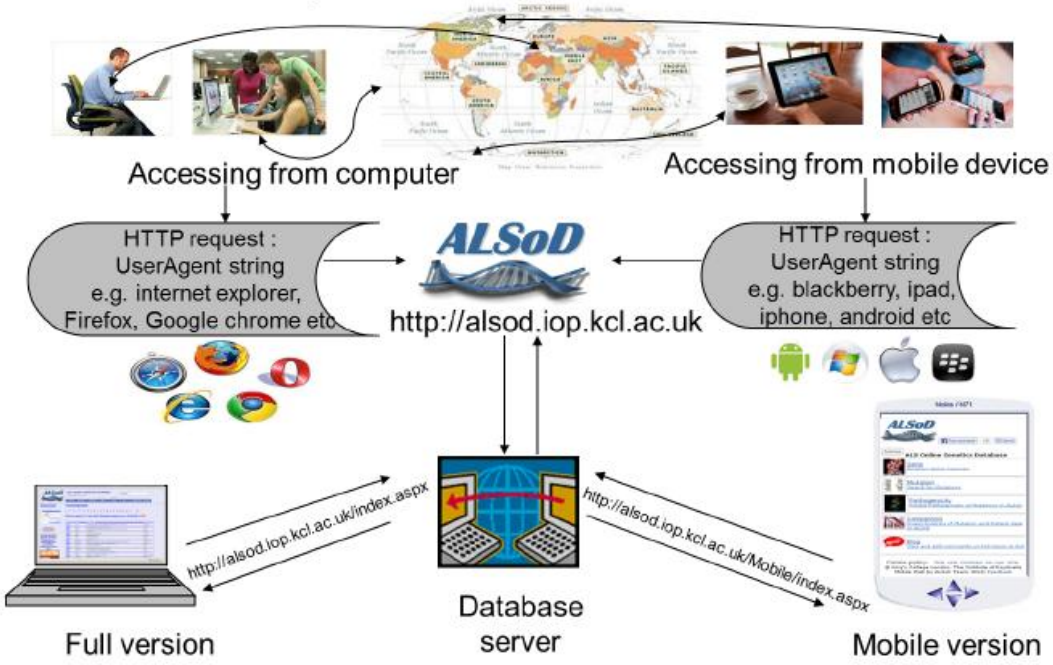
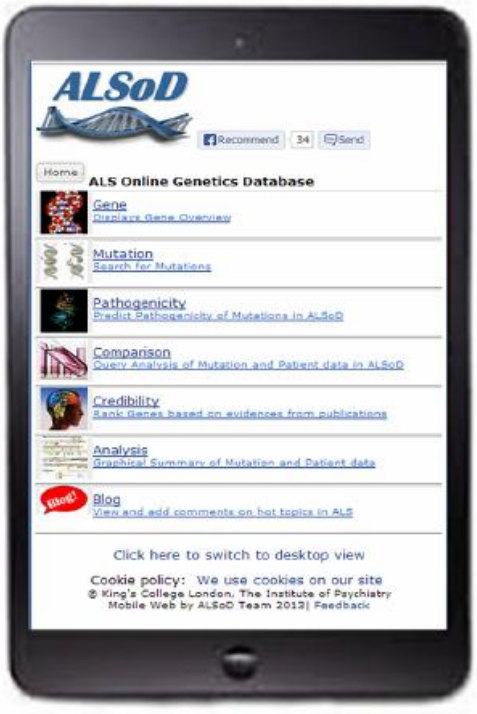


Figure 2. Mobile view of website.



App Development

The most commonly used smartphone platforms are iOS, BlackBerry, Android, and Windows Mobile. We used Eclipse software as the integrated development environment (IDE), with the Android software development kit (SDK), Android development tools (ADT) plugin, BlackBerry plugin, and the latest SDK tools and platforms, downloaded using the SDK Manager [18-22].

App Submission

From Eclipse, the application was compiled producing an .apk file (Android application package file format). This file was submitted to a registered Google Play account with a generated keystore containing a private key [23]. The ALSoD app can be downloaded from Google Play and currently, our Google app account confirms that the ALSoD app has had more than 100 downloads. The app then displays the website (designed using the Microsoft ASP.NET framework) with no status bar or URL navigation on the screen.

Creating Awareness

A marquee function scrolling text from right to left was inserted on the desktop master page to create an alert for regular users of the website, as seen in Figure 3. At symposiums and seminars, researchers were exposed to the recent development of the mobile app, which has contributed to increased Web traffic to ALSoD.

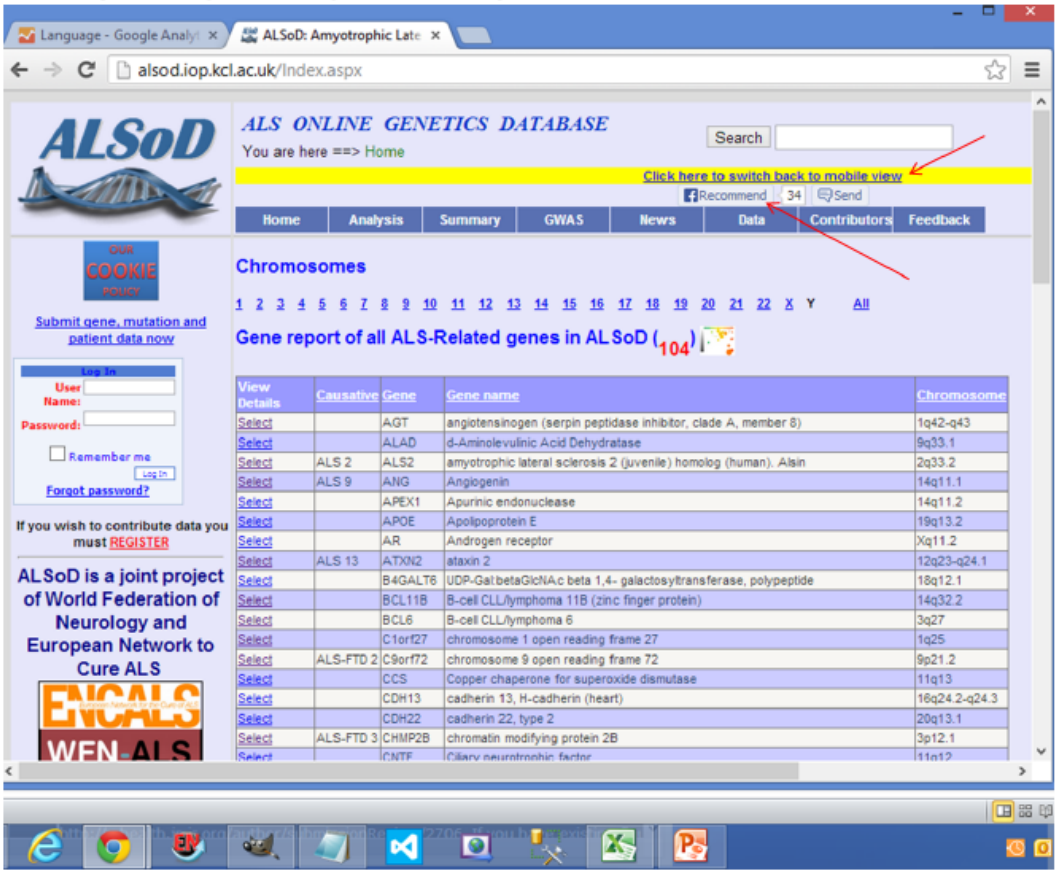
Feedback From App Users

During the various presentations of the mobile app development through posters and seminars, practical assessment of the website on mobile phones was carried out by attendants. Responses, questions asked, concerns raised, and critical analysis given by the audience were recorded and considered.

Analysis of Visits

The Google Analytics account for ALSoD was created in August 2012 to compare results generated by the two tools (mobile website and app) and to gain insight into the changes of the design and content of the website [24]. Visits were compared for the period before mobile website development, from August 2012 to October 2012, with the period after, from November 2012 to January 2013.

Figure 3. Desktop view creating awareness using Facebook and a marquee.



Results

Optimization of Webpages

In the 6-month data analysis period from August 2012 to January 2013, there were 5051 visits to the website, of which 2698 were unique (53%). There were 19,785 page views, of which 8883 (45%) were to 4 sets of pages. These pages focused on the pathogenicity of mutations, gene information, data analysis of mutations, and patient data.

Design Heuristics

Pages were optimized for mobile browsing by reducing image size to 5% of the original size, creating a mobile master page different from the desktop master page, and creating a link page to allow users to switch seamlessly between the mobile and desktop views.

Mobile Device Detection

If the UserAgent string contained keywords suggesting a mobile platform, for example, BlackBerry, Palm, mobile, iPhone, or iPad, then the user's device was redirected to the mobile site [25] displaying the compact version (Figure 2) of the website instead of the full version (Figure 3) [11].

Requesting Responses

To test this, we sent the mobile site URL to mobile phones of 14 users (colleagues and friends from whom we could easily obtain verbal feedback): 3 on the Android platform, 1 on the Windows Phone, 5 on BlackBerry OS, and 5 on iOS (2 iPhone and 3 iPad). All users gave positive feedback except for the Windows Phone user who could not utilize the pages with dropdown boxes.

App Development

Following successful implementation of the mobile website, we began app development. One straightforward method to achieve this is to automatically convert an already-built mobile website into a native app. This is done through the "WebView" object, which is an in-app Web browser used to display a website as if viewed on the browser of an Android smartphone [11]. For testing, we downloaded and used Android simulators. The plugins allow programmers to develop, test, and debug a Java application using the Eclipse IDE, but it requires a high level of programming skill [26]. We also tested and manipulated the .apk file on a real Android phone before submitting to Google Play.

Creating Awareness

Following the development of the mobile version, on the ALSoD Facebook page, 34 users recommended the website by May 29, 2013, using the Facebook "Recommend" button embedded on the website. Current tabular data are available on the website [27], displaying the growth of visits to the genetic database, as seen graphically in Figure 4.

Feedback From App Users

After the creation, testing, and publicity of the app, we received feedback from users about: caching for offline viewing [28-30], which would enable users to continue work; having a "page loading" icon when connecting; making users aware of the cookies policy; using an option menu button [31] to display analysis webpages (interaction.aspx, credibility.aspx, analysis.aspx); and creating a link to allow users to switch from mobile view to desktop view, as this would be useful on tablets like the iPad. We were able to implement all changes except for the offline viewing, which is difficult to implement because the database is large and held online.

Analysis of Visits

Our Google Analytics account showed that visits to the website increased from 2231 to 2820, yielding a 26% increase from pre-mobile period to post-mobile period and a 230% increase on the use of mobile devices (including tablets) to access the ALSoD website. On average, there were 300 unique visitors a day suggesting a high demand from the research community. A total of 1595 unique visitors in the post-mobile era accessed 11,376 page views on the website as opposed to 1220 unique visitors in the pre-mobile period (an increase of 31%), showing the relevance of a mobile-friendly website. Five mobile operating systems (Android, iOS, BlackBerry, Windows Phone, Symbian) were detected to have accessed the website within 6 months. Although BlackBerry OS visits declined from 34 to 14 visits (58%), iOS for iPhones and iPads increased from 40 to 105 visits (162%), and visits by Android devices increased from 29 to 213 visits (634%) (see Table 1). The Google search engine was the most used to search for the website (see Table 2). The likely explanation for the great increase in the use of Android devices is the development and introduction of the Android app submitted to Google Play.

Table 1. Comparison of website visits between the pre-mobile and post-mobile development.

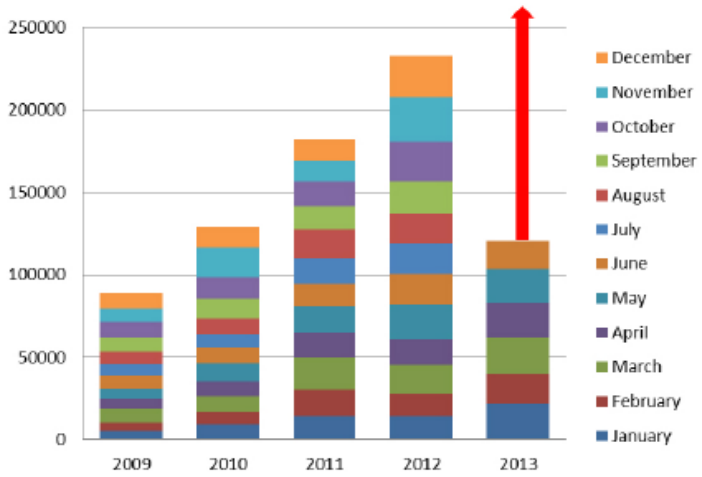
Operating system		Visits	Pages per visit	Avg visit duration	% new visits	Bounce rate, %
Totals						
		26.40%	7.03%	1.48%	4.25	2.98
		2820 vs 2231	4.03 vs 3.77	00:03:55 vs 00:03:58	52.45 vs 54.77	47.45 vs 48.90
Windows						
	01/Nov/2012-31/Jan/2013	1945	4.25	00:04:20	49.56	44.78
	01/Aug/2012-31/Oct/2012	1564	3.85	00:04:13	56.46	48.34
	% Change	24.36	10.47	2.71	-12.21	-7.36
Macintosh						
	01/Nov/2012-31/Jan/2013	435	3.26	00:02:36	51.26	53.33
	01/Aug/2012-31/Oct/2012	490	3.27	00:02:50	51.02	54.08
	% Change	-11.22	-0.28	-8.71	0.48	-1.38
Android						
	01/Nov/2012-31/Jan/2013	213	2.35	00:01:29	76.53	63.85
	01/Aug/2012-31/Oct/2012	29	4.93	00:03:12	65.52	44.83
	% Change	634.48	-52.30	-53.69	16.80	42.43
iOS						
	01/Nov/2012-31/Jan/2013	105	4.52	00:02:31	82.86	48.57
	01/Aug/2012-31/Oct/2012	40	4.55	00:02:06	87.50	50.00
	% Change	162.50	-0.58	19.77	-5.31	-2.86
Linux						
	01/Nov/2012-31/Jan/2013	85	6.51	00:08:11	28.24	29.41
	01/Aug/2012-31/Oct/2012	61	3.23	00:03:13	34.43	32.79
	% Change	39.34	101.45	154.08	-17.98	-10.29
Other systems						
	01/Nov/2012-31/Jan/2013	15	1.07	00:00:28	100.00	93.33
	01/Aug/2012-31/Oct/2012	13	1.92	00:00:43	92.31	84.62
	% Change	15.38	-44.53	-35.61	8.33	10.30
BlackBerry						
	01/Nov/2012-31/Jan/2013	14	8.14	00:11:10	0.00	28.57
	01/Aug/2012-31/Oct/2012	34	7.12	00:14:12	5.88	17.65
	% Change	-58.82	14.40	-21.31	-100.00	61.90
Windows Phone						
	01/Nov/2012-31/Jan/2013	4	6.75	00:08:02	50.00	50.00
	01/Aug/2012-31/Oct/2012	0	0	00:00:00	0.00	0.00
	% Change	∞	∞	∞	∞	∞
LG						
	01/Nov/2012-31/Jan/2013	3	1.33	00:00:11	0.00	66.67
	01/Aug/2012-31/Oct/2012	0	0	00:00:00	0.00	0.00
	% Change	∞	∞	∞	0.00	∞
Samsung						
	01/Nov/2012-31/Jan/2013	1	1	00:00:00	100.00	100.00

Operating system	Visits	Pages per visit	Avg visit duration	% new visits	Bounce rate, %
01/Aug/2012-31/Oct/2012	0	0	00:00:00	0.00	0.00
% Change	∞	∞	0.00	∞	∞

Table 2. Referral traffic from search engines from August 2012 to June 2013.

Source	Visits	Pages per visit	Avg. visit duration	% new visits	Bounce rate, %
Google	4339	4.06	00:04:54	42.73	46.58
Yahoo	62	3.63	00:02:42	51.61	41.94
Bing	32	5.53	00:07:11	62.50	31.25
Baidu	21	6.19	00:12:50	23.81	23.81
Daum	13	3.46	00:06:35	30.77	30.77
Ask	10	4.9	00:08:33	20.00	20.00
Conduit	8	17.62	00:08:27	37.50	25.00
AOL	7	1	00:00:00	71.43	100.00
Other search engines	5	6	00:04:24	0.00	0.00
Yandex	4	1.25	00:00:04	100.00	75.00
Babylon	3	1.33	00:00:33	100.00	66.67
AVG	1	1	00:00:00	100.00	100.00
Comcast	1	1	00:00:00	100.00	100.00

Figure 4. Graphical display of increased Web traffic on ALSoD.



Discussion

Principal Findings

ALSoD was developed as a disease-specific database for ALS, focused on genetics and phenotype, with planned incorporation of environmental and other risk factors in the future. We have shown that development of the website to facilitate smartphone access has greatly increased access.

In broad terms, there are two strategies for development of an app like this. Either the app can be developed as stand-alone software or it can be developed as a means to access an existing mobile website. We chose access to a mobile website for several reasons. First, we used this approach because mobile websites are immediately accessible to users through a browser, more compatible across devices, have easier content updates, are faster to find on search engines, make it easier to share content via a link, have a longer lifecycle on a user's device, are easily convertible to an app, and are more cost-effective [14,32]. Second, the database is regularly updated with data, which would require the release of weekly updates to an app if the website was not the primary content holder. Third, although third-party automated app development tools exist [33,34], it

was simple for us to convert the mobile website into an Android app using the WebView object.

A limitation to this approach is that the website recognizes a mobile device based on the information held in the list in the UserAgents string. Although software exists to automatically update the list, there is a financial cost involved and we have therefore chosen to update the list manually. Furthermore, UserAgents strings are limited in the kind of information sent to the server. Specific information such as the size of the screen, the manufacturer, the format of image supported, and the model of phone are not sent. This is an issue because some mobile devices allow portrait and landscape views when repositioned while others have unique width and height dimensions. It is therefore difficult to have a perfect display on all devices.

There are 3 main methods by which users access the website (Table 2). Roughly 39% of the traffic is organic from the Google search engine, 37% is direct by typing the ALSoD URL directly on a browser, and the rest are referrals through external sites collaborating with ALSoD.

More than 100 interactive, downloadable widgets and mobile applications have been submitted to the NHS Choices Health Apps library [35]. Some of these apps are commercial apps and the freely available ones range from calculating alcohol consumption to weight tracking. There are no specific genetics disease apps that concentrate on combining genotype, phenotype, and geographical information with associated analysis tools, although ALS database apps do exist. For example, the PatientsLikeMe app was initially developed to help United Kingdom-based ALS patients find clinical trials that are right for them and organizations find patients right for their trials [36,37]. The ALSoD app has direct relevance to clinicians working in ALS and therefore relevance to the NHS Choices Health Apps library. It includes a comparison tool to evaluate information for different genes side by side or jointly with user configurable features, a pathogenicity prediction tool using a combination of computational approaches to distinguish variants with nonfunctional characteristics from disease-associated mutations with more dangerous consequences, and a credibility tool to enable ALS clinicians and researchers to objectively assess the evidence for gene-causation in ALS. A checklist, as seen in Multimedia Appendix 2, was used to report the web-based intervention of users [38]. Furthermore, integration of external tools, systems for feedback, annotation by users, and two-way links to collaborators hosting complementary databases further enhance the functionality.

Conclusions

Development of the mobile website and associated app has increased access to this disease-specific database and facilitated access through a wide range of devices. Visitor analysis has shown the importance of collaborating with other relevant databases through hyperlinks. Our future work will concentrate on further integration with other databases, adding in nongenetic risk factors, and increasing access and relevance for related research disciplines.

Acknowledgments

We are especially grateful for the long-standing and continued funding of this project from the ALS Association and the MND Association of Great Britain and Northern Ireland. We also thank ALS Canada, MNDA Iceland, and the ALS Therapy Alliance for support. The research leading to these results has received funding from the European Community's Health Seventh Framework Programme FP7/2007-2013, under grant agreement number 259867. We thank the NIHR specialist Biomedical Research Centre for Mental Health and the Biomedical Research Unit for Dementia at the South London and Maudsley NHS Foundation Trust (SLaM) and the Institute of Psychiatry, King's College London. Aleks Radunovic, Nigel Leigh, and Ian Gowrie originally conceived ALSoD. ALSoD is a joint project of the World Federation of Neurology (WFN) and European Network for the Cure of ALS (ENCALS).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Script for redirecting views.

[PDF File (Adobe PDF File), 20KB - mhealth_v1i2e18_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.2 [38].

[PDF File (Adobe PDF File), 982KB - mhealth_v1i2e18_app2.pdf]

References

1. Al-Chalabi A, Jones A, Troakes C, King A, Al-Sarraj S, van den Berg LH. The genetics and neuropathology of amyotrophic lateral sclerosis. Acta Neuropathol 2012 Sep;124(3):339-352. [doi: 10.1007/s00401-012-1022-4] [Medline: 22903397]
2. ALSoD Website. URL: http://alsod.iop.kcl.ac.uk/ [accessed 2013-07-26] [WebCite Cache ID 6IOt0AEI6]

3. Abel O, Powell JF, Andersen PM, Al-Chalabi A. ALSoD: A user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics. Hum Mutat 2012 Sep;33(9):1345-1351. [doi: 10.1002/humu.22157] [Medline: 22753137]
4. Lill CM, Abel O, Bertram L, Al-Chalabi A. Keeping up with genetic discoveries in amyotrophic lateral sclerosis: the ALSoD and ALSGene databases. Amyotroph Lateral Scler 2011 Jul;12(4):238-249. [doi: 10.3109/17482968.2011.584629] [Medline: 21702733]
5. Bickmore TW, Schilit BN. Digestor: device-independent access to the World Wide Web. Computer Networks and ISDN Systems 1997 Sep;29(8-13):1075-1082. [doi: 10.1016/S0169-7552(97)00026-3]
6. Cutshall T, Blake L, Bandy S. Creating a Mobile Library Website. 2011. URL: http://gru.openrepository.com/gru/bitstream/10675.2/431/3/65258930.pdf [accessed 2013-05-01] [WebCite Cache ID 6GIupVdnJ]
7. Gube J. Mashable Business. 2010. 8 Tools For Easily Creating a Mobile Version of Your Website URL: http://mashable.com/2010/12/16/create-mobile-site-tools/ [accessed 2013-05-01] [WebCite Cache ID 6GIuG7UJE]
8. Bosomworth D. Smart Insights. 2012. Mobile Marketing Statistics 2012 URL: http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/ [accessed 2013-05-01] [WebCite Cache ID 6GIu6J90i]
9. Maurer ME, Hausen D, De Luca A, Hussmann H, editors. Mobile or desktop websites?: website usage on multitouch devices. ACM 2010. [doi: 10.1145/1868914.1869018]
10. Agar J. Constant Touch: A Global History of the Mobile Phone. Cambridge: Icon Books; 2004.
11. Iglesias E, Meesangnil W. Mobile website development: From site to app. Bul Am Soc Info Sci Tech 2011 Dec 01;38(1):18-23. [doi: 10.1002/bult.2011.1720380108]
12. Imielinski T, Korth H. Introduction to mobile computing. Mobile Computing 1996:1-43. [doi: 10.1007/978-0-585-29603-6_1]
13. Lemaire ED, Boudrias Y, Greene G. Low-bandwidth, Internet-based videoconferencing for physical rehabilitation consultations. Journal of Telemedicine and Telecare 2001 Apr 01;7(2):82-89. [doi: 10.1258/1357633011936200]
14. Summerfield J. Mobile Website vs Mobile App (Application): Which is Best for Your Organization?.: HumanServiceSolutions; 2012. URL: http://www.hswsolutions.com/services/mobile-web-development/mobile-website-vs-apps/ [accessed 2013-05-01] [WebCite Cache ID 6GIvmP4Lv]
15. Flow A. How to Make a Mobile Website - 6 Easy Tips. 2010. URL: http://blog.templatemonster.com/2010/05/11/how-make-mobile-website-6-easy-tips/ [accessed 2013-05-01] [WebCite Cache ID 6GIvv30D8]
16. Georgiou N. ASP.NET IsMobileDevice and how it works. 2011 Mar 12. URL: http://stackoverflow.com/questions/1829089/how-does-ismobiledevice-work [accessed 2013-05-01] [WebCite Cache ID 6GJ1Pw4DC]
17. Ellison NB, Steinfield C, Lampe C. The benefits of Facebook "friends": Social capital and college students' use of online social network sites. Journal of Computer-Mediated Communication 2007;12(4):1143-1168. [doi: 10.1111/j.1083-6101.2007.00367.x]
18. Android Developers. Building Your First App. 2012. URL: http://developer.android.com/training/basics/firstapp/creating-project.html [accessed 2013-05-01] [WebCite Cache ID 6GIx28QJm]
19. Android Developers. Web Apps. 2012. URL: http://developer.android.com/guide/webapps/webview.html [accessed 2013-05-01] [WebCite Cache ID 6GIxJIBSE]
20. Heuwieser FM. BlackberryCool. 2009. How to write your first BlackBerry application – Lesson 1: UI fields URL: http://www.blackberrycool.com/2009/06/05/how-to-write-your-first-blackberry-application-lesson-1-ui-fields/ [accessed 2012-10-01] [WebCite Cache ID 6GIxXzwd3v]
21. Genuitec. Genuitec Product and Eclipse Package Downloads. 2012. URL: http://www.genuitec.com/all_eclipse.html [accessed 2013-05-01] [WebCite Cache ID 6GIxn9tTR]
22. Blackberry Developer. BlackBerry Java 7.1 SDK. 2012. URL: https://developer.blackberry.com/java/download/ [accessed 2013-05-01] [WebCite Cache ID 6GIy1zAj5]
23. Google App Engine.: Developers G; 2012. URL: https://developers.google.com/appengine/docs/whatisgoogleappengine [accessed 2013-05-01] [WebCite Cache ID 6GIyRwppP]
24. Plaza B. Google Analytics for measuring website performance. Tourism Management 2011;32(3):477-481.
25. ALSoD Mobile Site. URL: http://alsod.iop.kcl.ac.uk/Mobile/index.aspx [accessed 2013-07-26] [WebCite Cache ID 6IOtI9Unv]
26. Greenfield M. CNET UK. 2012. How to develop Android apps URL: http://reviews.cnet.co.uk/mobile-apps/how-to-develop-android-apps-50008626/ [accessed 2013-07-26] [WebCite Cache ID 6IOrsFF1O]
27. ALSoD Visitor Statistics. URL: http://alsod.iop.kcl.ac.uk/charts/index.aspx [WebCite Cache]
28. Hickson I. World Wide Web Consortium, Working Draft WD-html5-20120329. 2012. A vocabulary and associated APIs for HTML and XHTML URL: http://www.w3.org/TR/2012/WD-html5-20120329/ [accessed 2013-05-01] [WebCite Cache ID 6GIz04ayC]
29. vandenbossche J. WINK(Webapp INnovation Kit). HTML5 Application Cache Study URL: http://www.winktoolkit.org/blog/235/ [accessed 2013-07-26] [WebCite Cache ID 6IOs9zfYo]
30. Grove R. Mobile Browser Cache Limits: Android, iOS, and webOS. 2010. URL: http://www.yuiblog.com/blog/2010/06/28/mobile-browser-cache-limits/ [accessed 2013-05-01] [WebCite Cache ID 6GIzC67wM]

31. Smith S. Android SDK: Implement an Options Menu. 2012. URL: http://mobile.tutsplus.com/tutorials/android/android-sdk-implement-an-options-menu/ [accessed 2013-07-26] [WebCite Cache ID 6IOsIvgNo]
32. Klein D. How to decide: Mobile websites vs mobile apps. 2013. URL: http://www.adobe.com/inspire/2012/02/mobile-websites-vs-mobile-apps.html [accessed 2013-05-01] [WebCite Cache ID 6GIzmzWi6]
33. mobincube. 2012. Mobimento Mobile SL URL: http://www.mobincube.com/ [accessed 2013-05-01] [WebCite Cache ID 6GIzsJyqp]
34. AppsGeyser. 2012. URL: http://www.appsgeyser.com/ [accessed 2013-05-01] [WebCite Cache ID 6GIzumb3J]
35. NHS Choices Annual Report. 2011. NHS Choices. Improving Health, Improving Lives URL: http://www.nhs.uk/aboutNHSChoices/professionals/developments/Documents/annual-report/Annual_report_2011_digital.pdf [accessed 2013-07-31] [WebCite Cache ID 6IWnLVAab]
36. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. J Med Internet Res 2010;12(2):e19 [FREE Full text] [doi: 10.2196/jmir.1549] [Medline: 20542858]
37. PatientsLikeMe. PatientsLikeMe app. 2013. URL: http://apps.nhs.uk/app/patientslikeme/ [accessed 2013-05-01] [WebCite Cache ID 6GJ0Ftsae]
38. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]

Abbreviations

ALSoD: Amyotrophic Lateral Sclerosis Online Genetics Database
ASP.NET: active server page network
IDE: integrated development environment
Master page: provides automatic layout, pagination, headers and footers, and graphic elements for multiple pages.
OS: operating system
SMS: short message service
UserAgent: a user agent is software (a software agent) that is acting on behalf of a user.
Wi-Fi: wireless local area network

Edited by G Eysenbach; submitted 13.05.13; peer-reviewed by P Wicks, S Singh; comments to author 11.06.13; revised version received 21.06.13; accepted 30.06.13; published 04.09.13

Please cite as:
Abel O, Shatunov A, Jones AR, Andersen PM, Powell JF, Al-Chalabi A
Development of a Smartphone App for a Genetics Website: The Amyotrophic Lateral Sclerosis Online Genetics Database (ALSoD)
JMIR Mhealth Uhealth 2013;1(2):e18
URL: http://mhealth.jmir.org/2013/2/e18/
doi:10.2196/mhealth.2706
PMID:

Chapter 6 METHODS

6.1 Research questions with an overview of proposed methods

Description of the novel research questions proposed for my thesis project is:

1) Is it possible to generate a database that summarizes genetic data for a disease and allow meta-analysis online? Amyotrophic lateral sclerosis (ALS) will be used as a model disease.

I will create gene and chromosomal overview pages of all available ALS-related genes; collate published genome-wide association study (GWAS) data from the largest genome-wide association study of ALS to date completed by our group; collate published linkage studies and merge results from GWAS; and allow user-defined queries of meta analysis of both association and linkage studies.

2) Can all genetic data useful to ALS be automatically collected?

I will implement data mining and then manual curation of genetic data used by most databases; a tool to interrogate larger databases like UNIPROT (which is a consortium between the European Bioinformatics Institute (EBI), the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR)) will be developed.

3) How achievable is an on-the-fly analysis of genetic data online?

A tool for dynamically uploading GWAS data from users will be developed; I will compare user data with available GWAS data in the database, including appropriate statistical corrections and allow user-configurable queries using a simple interface.

4) To what extent can the database (the ALS Online Genetics Database) be integrated with other bioinformatics resources available online?

To provide users of the database with maximal cross-referenced bioinformatics investigation, I hope to utilize information from other well-known databases like Entrez Gene, UCSC, OMIM, Genecards, KEGG, Uniprot, iHop, dbSNP, Pubmed, Wolfram Alpha, PDB, Gene Onology and others.

5) With the advent of various ALS-related genes, can a database generate levels of evidence to support or refute genetic associations and linkages with ALS?

Various criteria were intended to generated scores for prioritizing the importance of each gene. The initially proposed criteria for each gene are: A score for the number of mutations related to ALS on ALSoD; a score for

the number of published articles related to ALS available in Pubmed; a score for the number of significant SNPs within a threshold derived from the meta-analysis tool; a score for how each gene relates to other genes using Ingenuity pathway tool; a score for the number of countries where mutations or disease-associated variants are found.

6.2 Keeping up with genetic discoveries in amyotrophic lateral sclerosis:

6.2.1 Data Submission

ALSoD allows users to submit new gene, mutation and patient data from publications or unpublished research work as displayed in Appendix 1 – Data Submission Flowchart. This is the only time a user requires a username and password to log in. A new submitter is required to register by providing details which will be stored on the database while an existing submitter logs in with login details already available on the database. Log in with a username and password either directly on the homepage or by clicking on the 'login' hyperlink to take you to the login page.

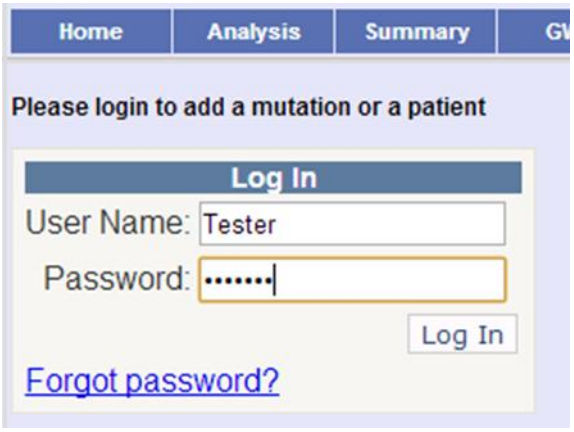


Figure 2: Logging in section

The protected page displaying a list of tasks that could be carried out by the secure user is shown in Figure 3.

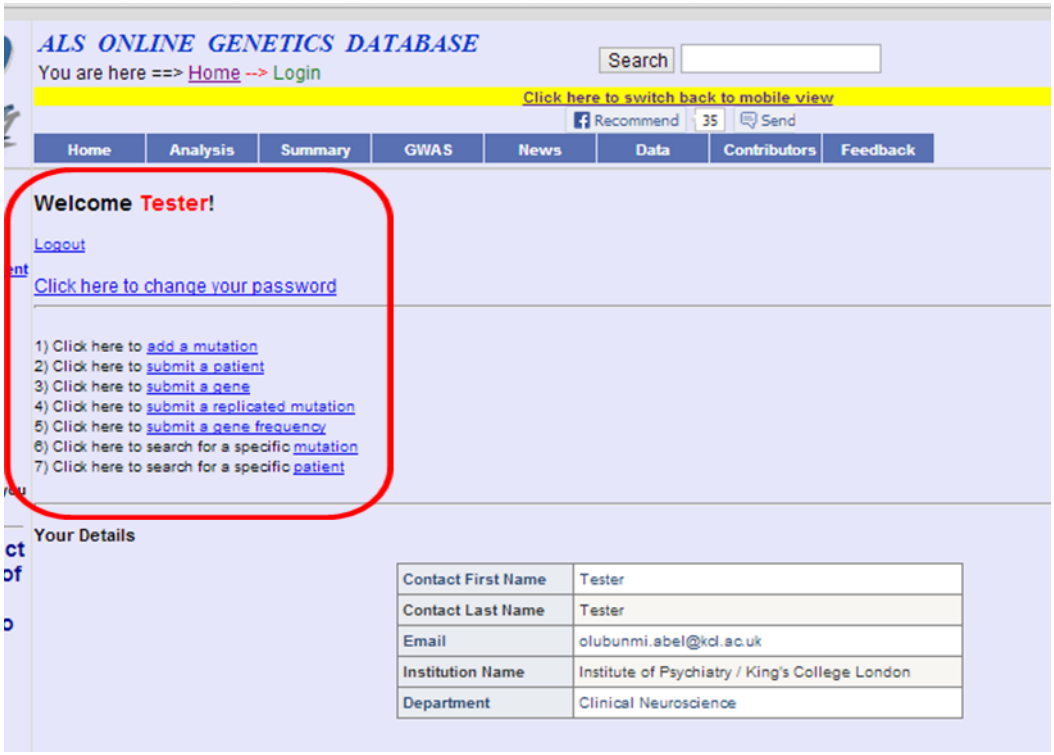


Figure 3: Protected page after log in

At the end of any successful submission to the database, a confirmation page appears as seen in Figure 4.

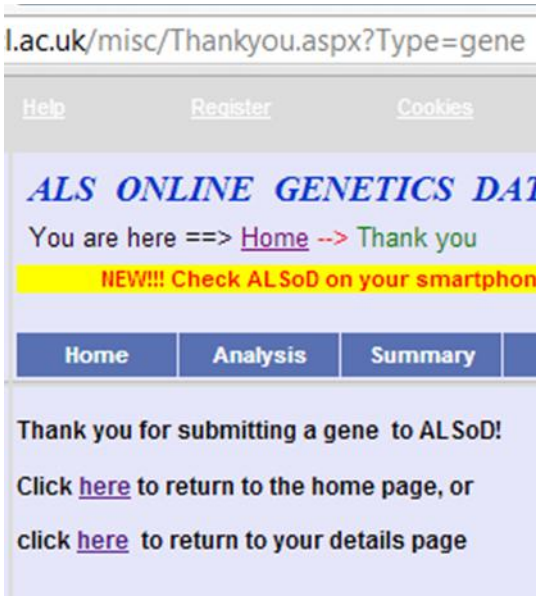


Figure 4: Confirmation page

6.2.2 Submit mutation data

6.2.2.1 SQL Script generating mutation table

A mutation table shown in Figure 5 was created to store data like codon, sequence position, gene, documentation and publication details on variants of genes recorded in publications or submitted by authentic researchers.

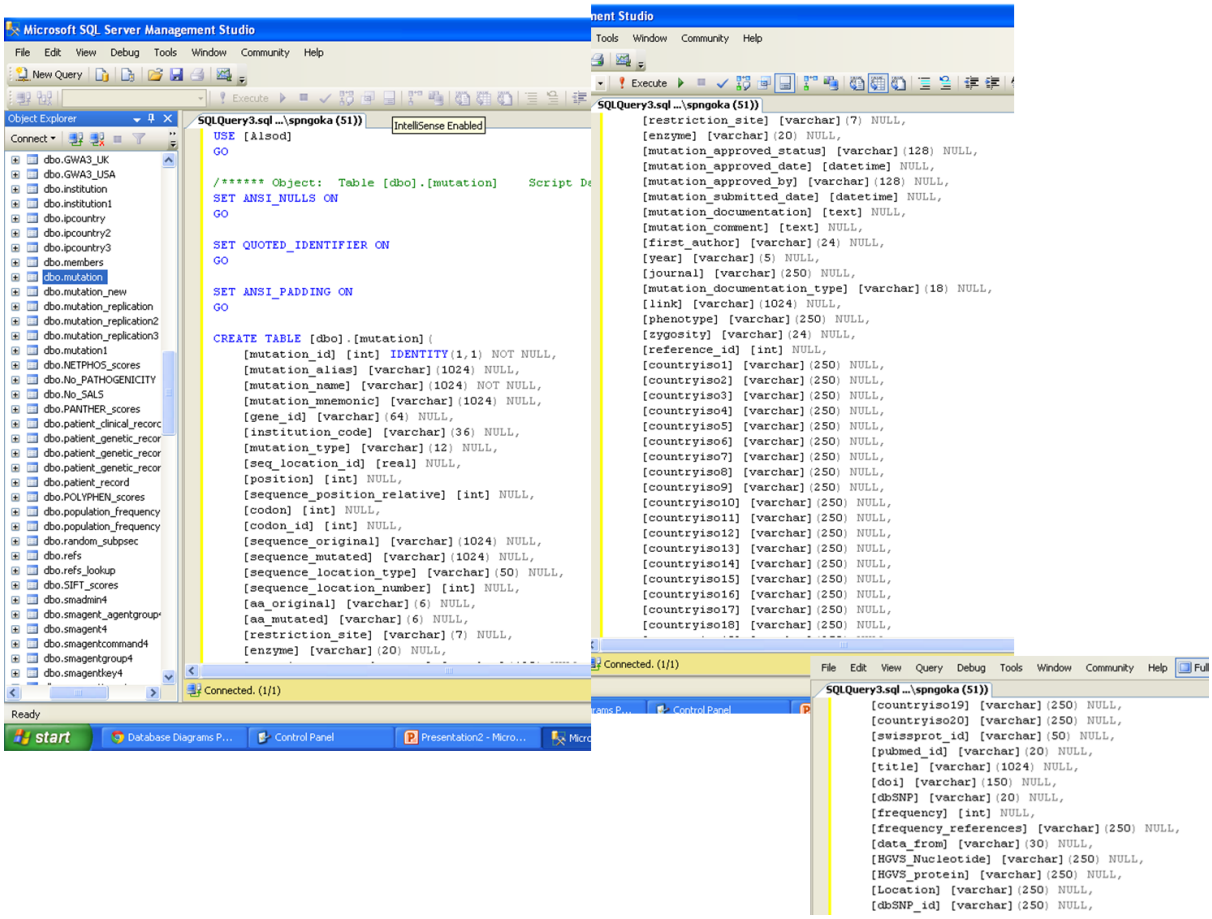


Figure 5: Mutation table on SQL Server

6.2.2.2 Submission process into ALSoD

Click on option 3 saying “click here to submit a gene” and it takes you to a list of genes to choose from. When I took over the database, there were only 4 genes (SOD1, ALS2, NEFH, and VAPB) available for mutation submission. Additional genes coded to the page are for TARDBP, ANG, FUS, SETX and OPTN as shown in Figure 6. More development for gene submission is still under construction.

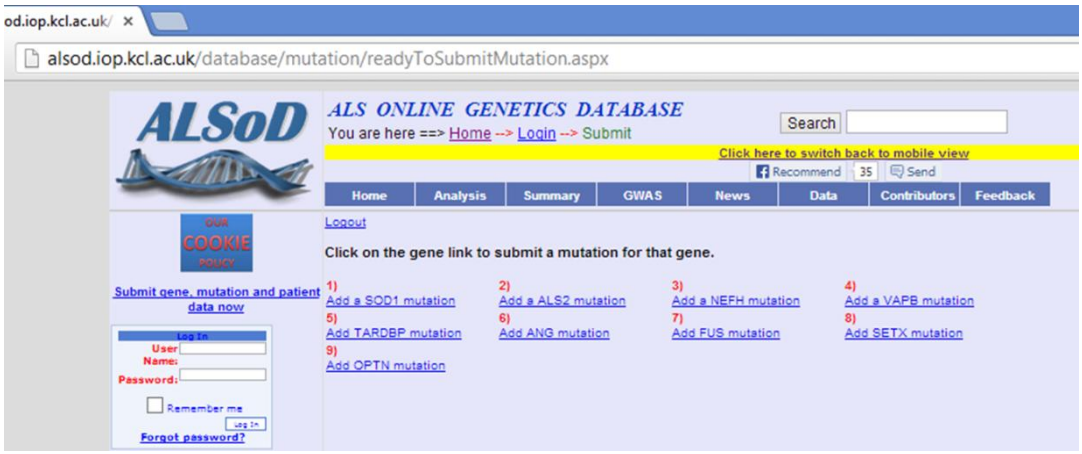


Figure 6: Gene submission page 1

Enter a mutation location (Exon or Intron) and type (Substitution, Insertion, Deletion, Polymorphism, Silent or Repeat) seen in Figure 7.

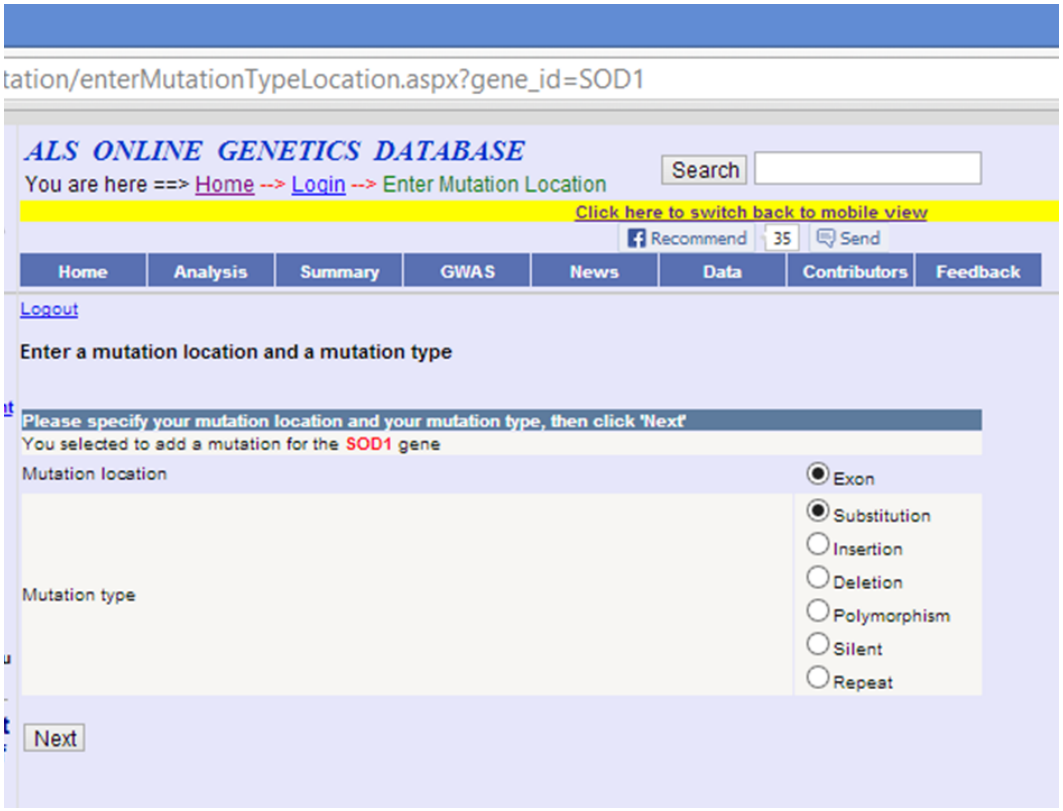


Figure 7: Gene submission page 2

Select the codon for the new mutation in Figure 8.

Figure 8: Gene submission page 3

Select the position of the trinucleotide mutated in Figure 9.

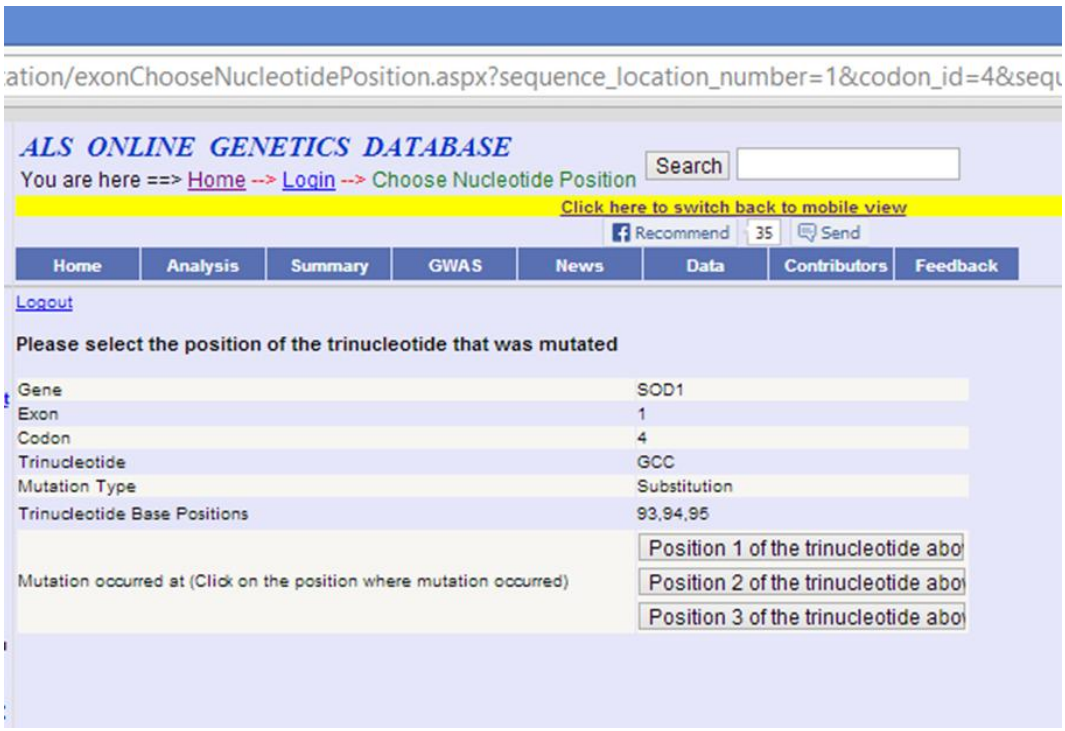


Figure 9: Gene submission page 4

Select the new amino produced by the mutation in Figure 10.

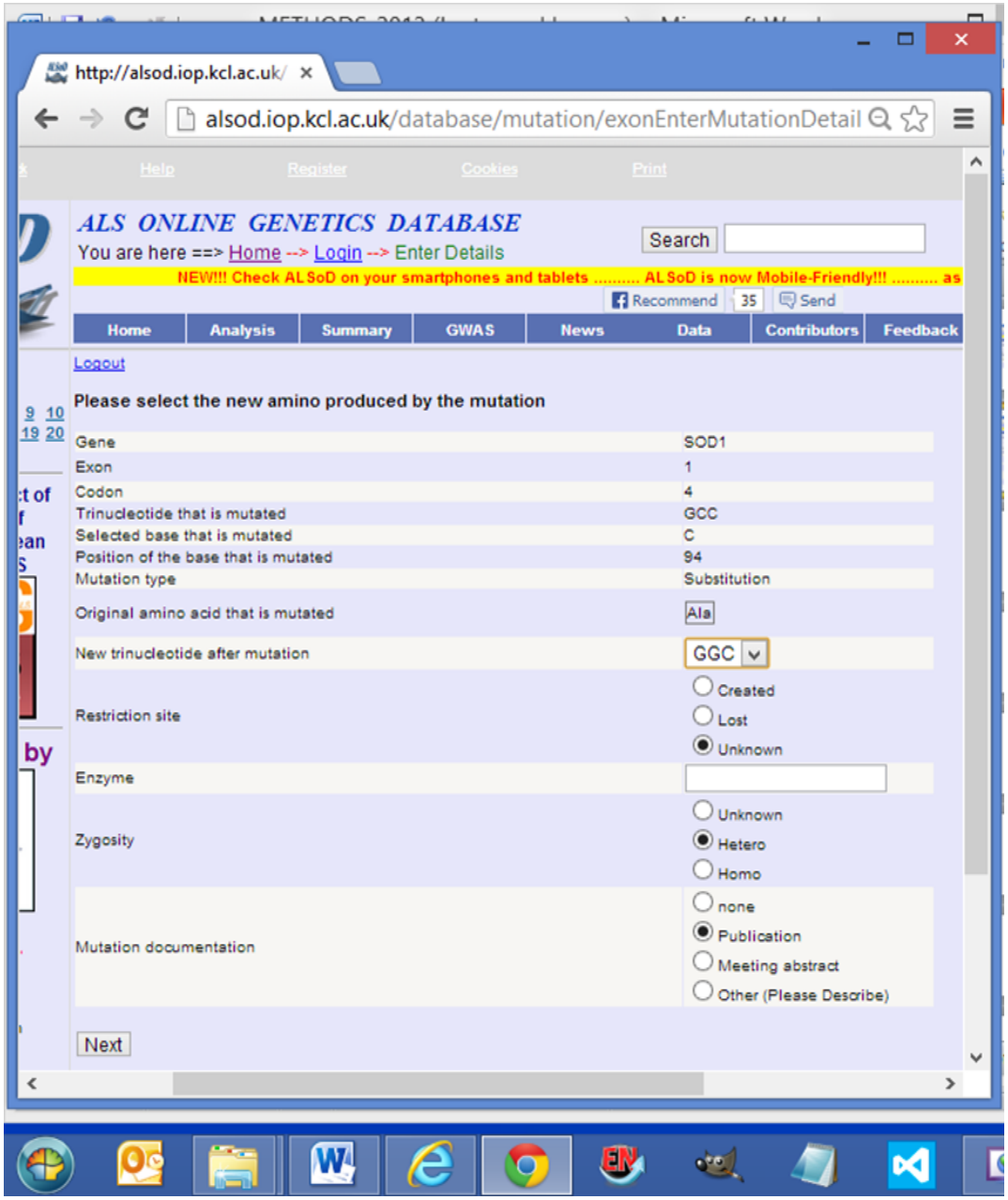


Figure 10: Gene submission page 5

Enter publication references of the mutation in Figure 11.

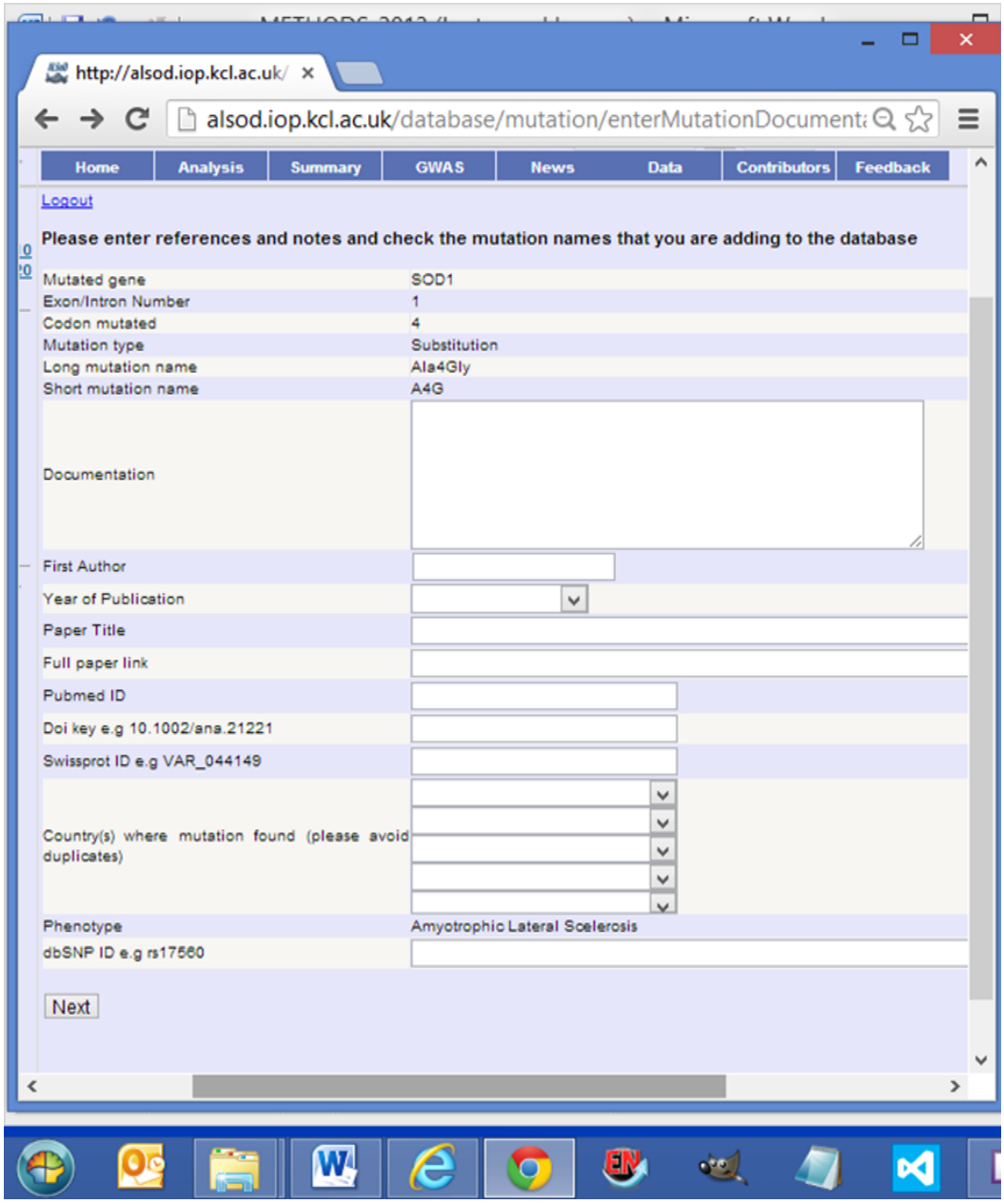


Figure 11: Gene submission page 6

Verify mutation details and submit in Figure 12.

Figure 12: Gene submission page 7

6.2.3 Submit new gene

6.2.3.1 SQL Script generating gene table

A gene table was created to store data like ids of other databases, chromosomal position and documentation details on genes recorded in publications shown in Figure 13.

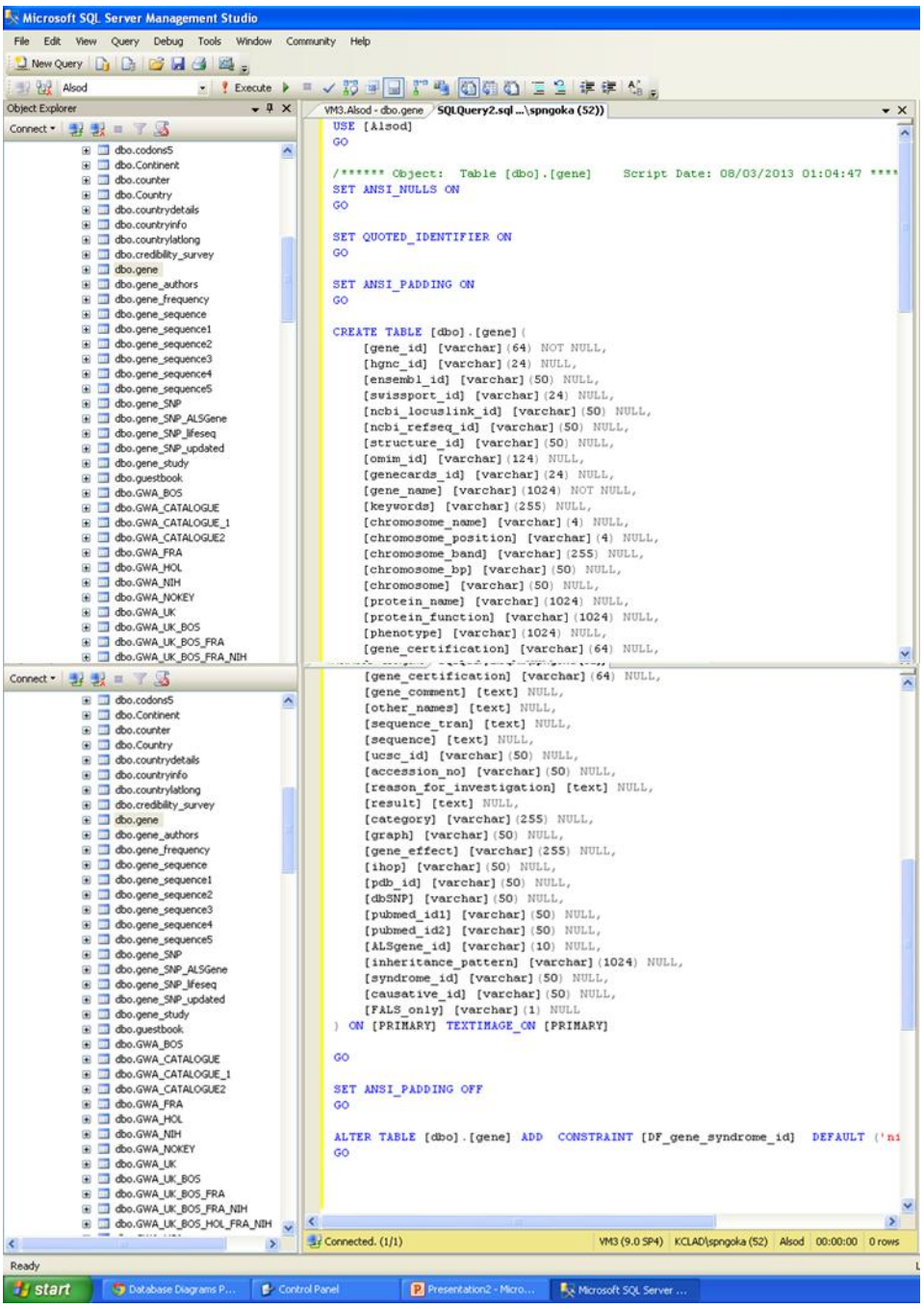


Figure 13: Gene table on SQL server

6.2.3.2 Submission process into ALSoD

I designed a single page for entering data required as a format for the gene table. The data are unique identity number for each gene in various databases to allow for database integration with other freely available databases related to specific genes shown in Figure 14. Most of these ids are extracted or curated from NCBI Entrez Gene, HUGO and iHop websites.

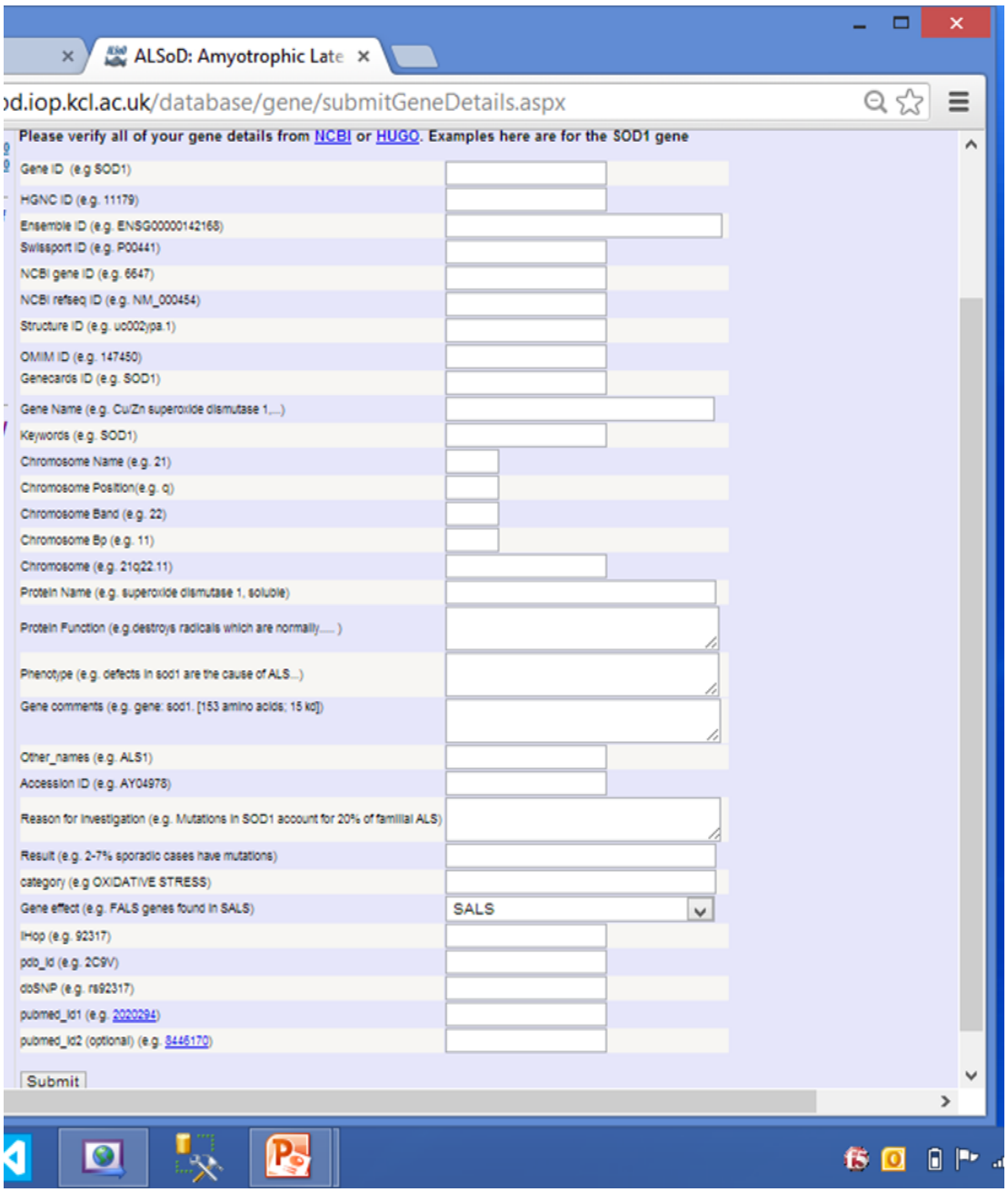


Figure 14: Gene submission page on ALSoD

Click on submit button and the data are automatically inserted into the gene table of the SQL ALSoD database.

6.2.4 Submit patient data

6.2.4.1 SQL Script generating patient table

A patient table was created to store data like age of onset, duration of the disease, site of onset, mutation and documentation details on patients with variants recorded in publications or submitted by authentic researchers as seen in Figure 15.

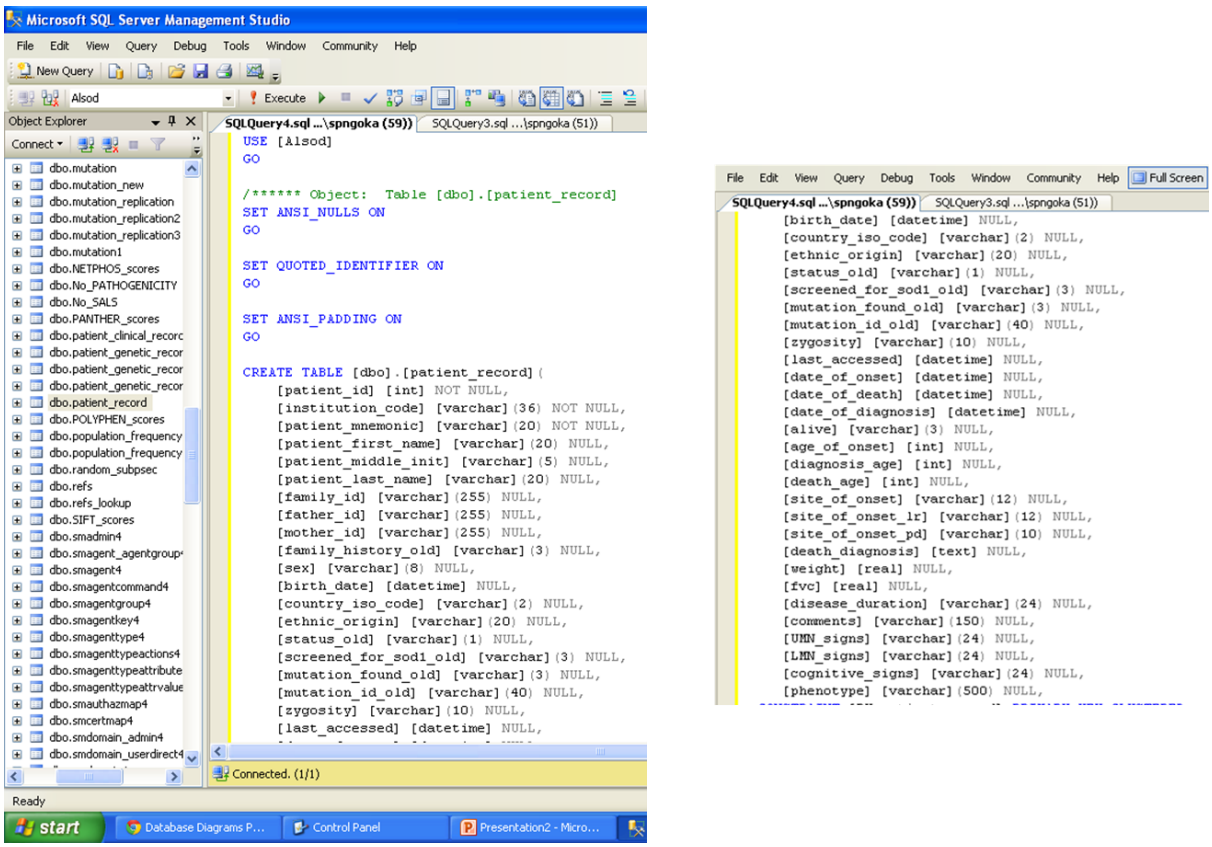


Figure 15: Patient table in SQL server

6.2.4.2 Submission process into ALSoD

A check is carried out to find out if the patient data needs to be submitted by asking the submitter some preliminary questions about the subjects like “Is the subject affected?” or “Does the subject have a mutation?” seen in Figure 16. If the answer to any of the questions is a “no”, the submission process is terminated but if “yes”, the process continues to the next page.

Figure 16: Patient submission page 1

The user is asked if the submission is for a new family or an old family. Before I took over, this field was used to allocate a unique family identity number to a family seen in Figure 17. This was however changed to the name of the first author with the year e.g. Al_Chalabi 1999. This was changed because not all patients are familial and it was my way of knowing the patient data submitted from a particular publication. Also, it is now illegal to disclose the identity of a patient [524] which could be traced from the family id. Confidentiality of patients is crucial in research work because it is illegal to disclose a patient's identity.

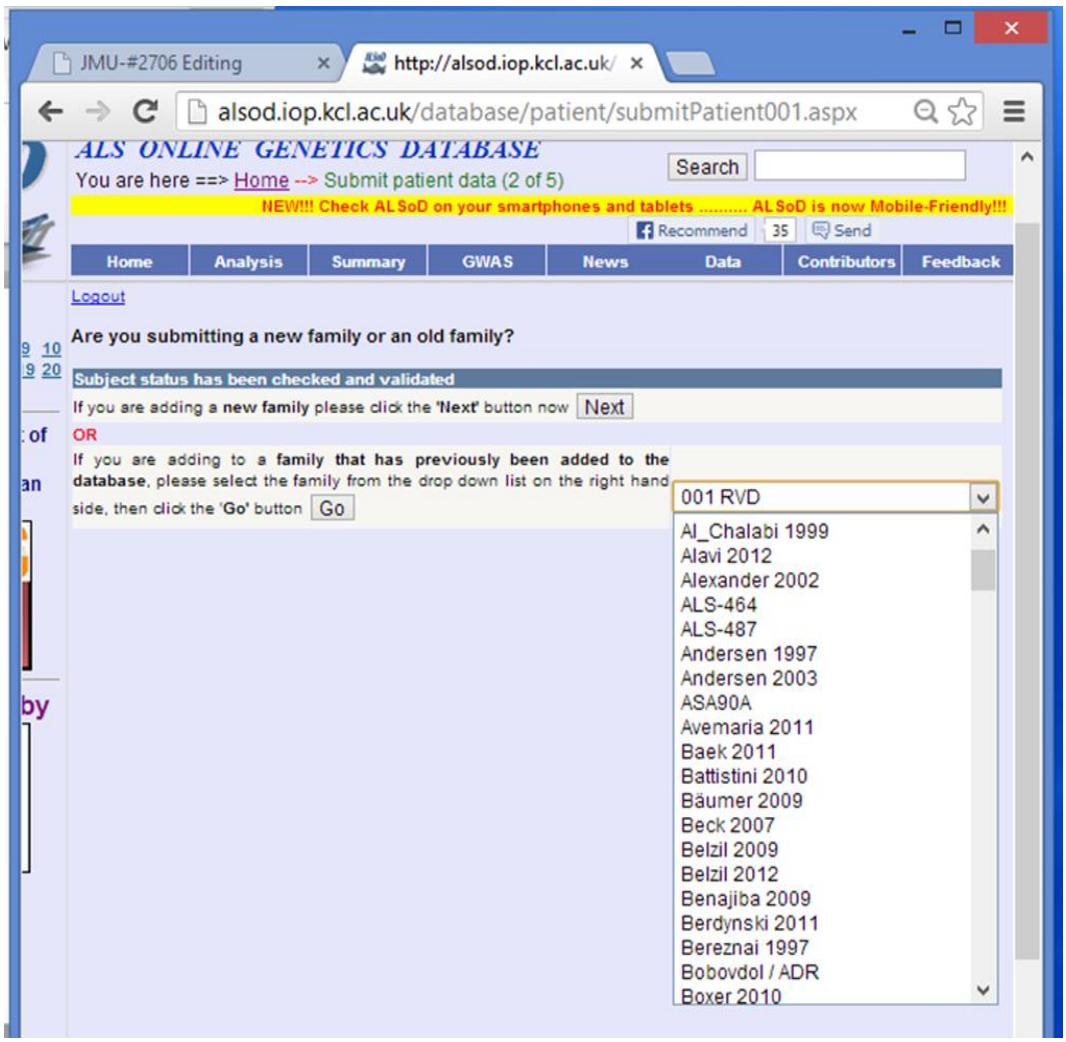


Figure 17: Patient submission page 2

Patient details like the gene, gender, family history are filled in. Some fields like the date of birth and family id were compulsory fields and the submitter could not move forward if the data was not provided. This has since been changed and no field is compulsory. The dropdownlist for a patient's country of origin and ethnicity with all the genes on the database are now included on this page in Figure 18.

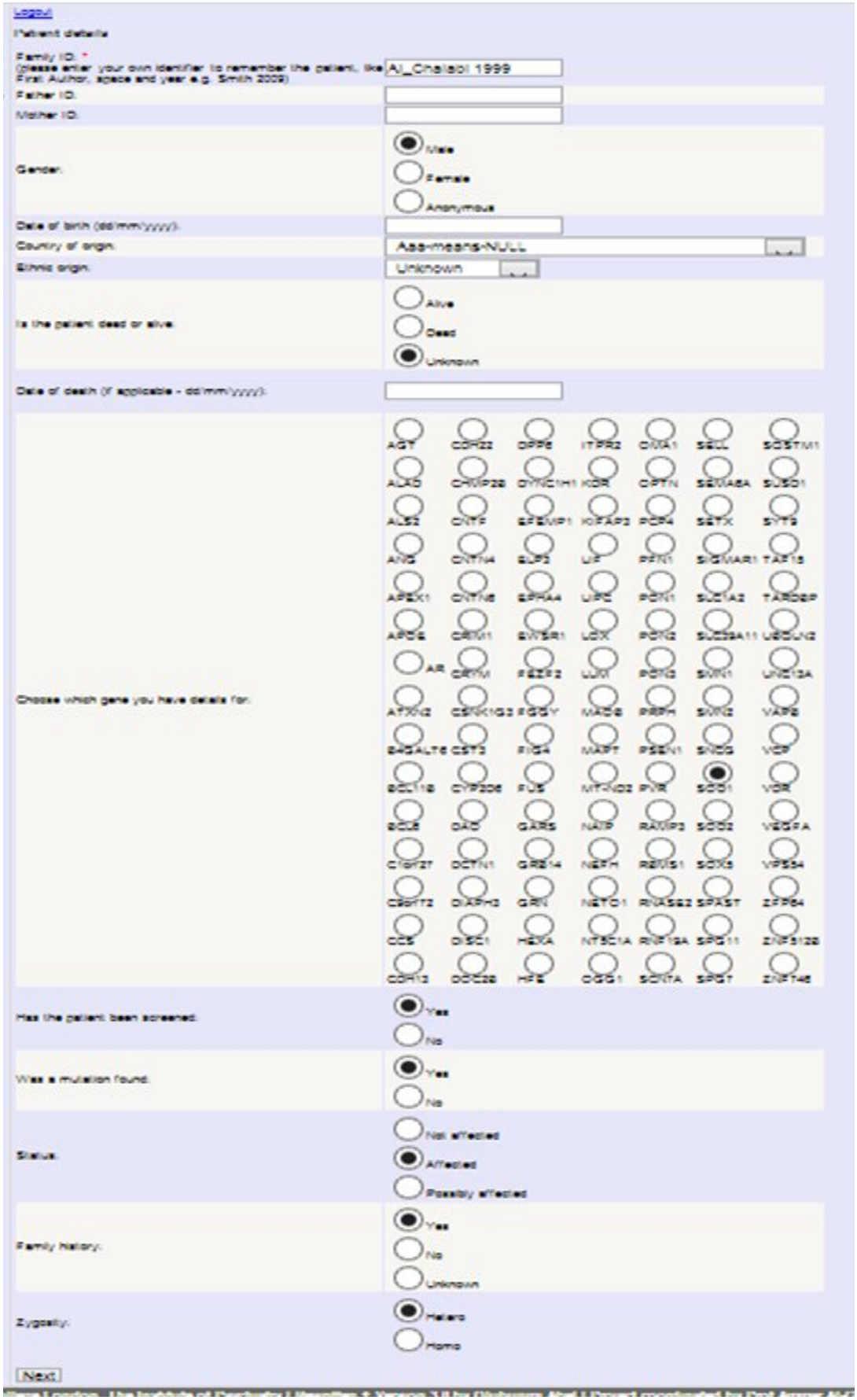


Figure 18: Patient submission page 3

Onset details like the site of onset, age of onset (in years) and duration (in months) are required here. Some of the fields are not important anymore and some like the site of onset was modified and updated as seen in Figure 19.

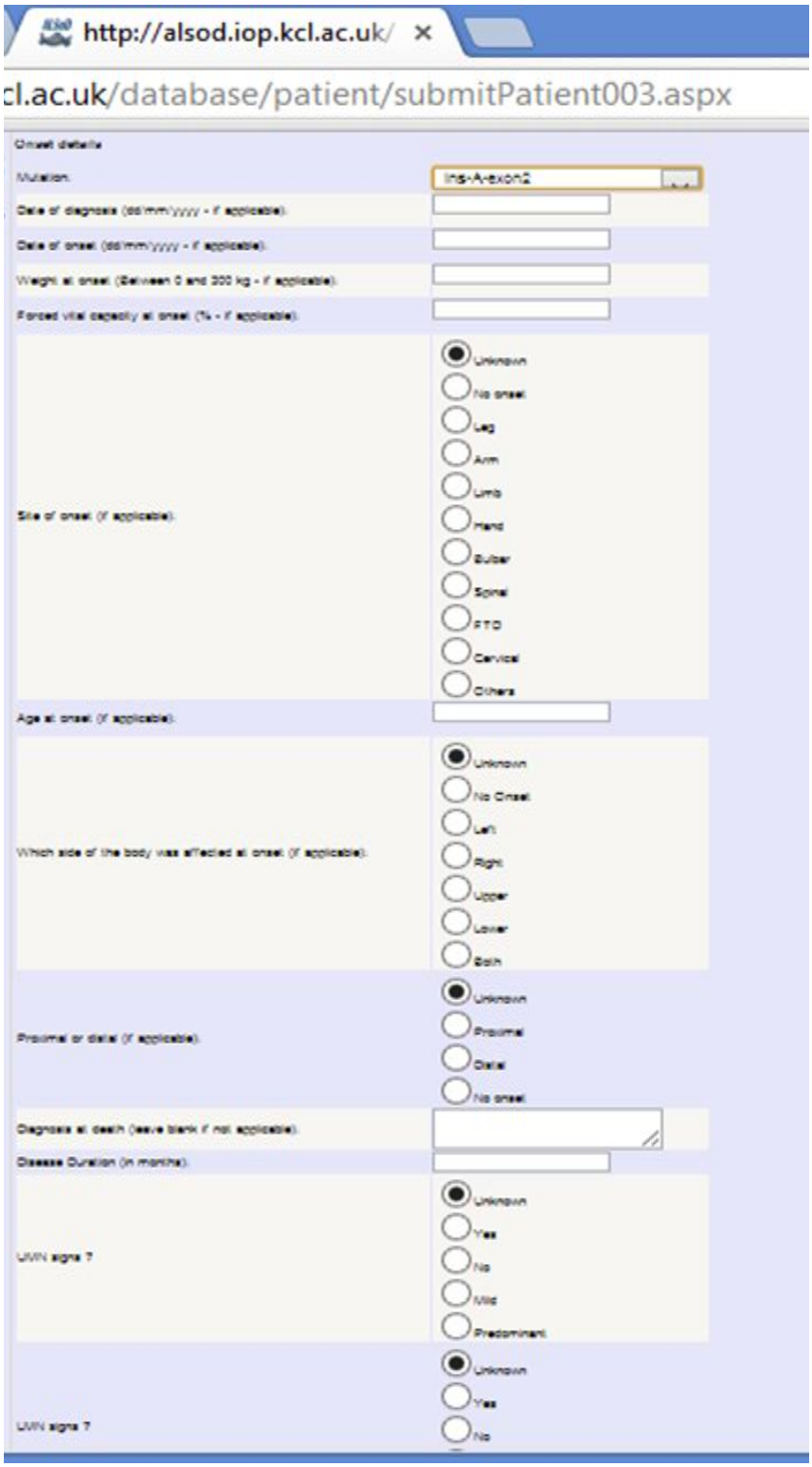


Figure 19: Patient submission page 4

Check the details inputted and click on ‘Submit’ button shown in Figure 20.

Figure 20: Patient submission page 5

The mutation data are inserted into the patient record, patient clinical records and the patient genetic record tables.

6.2.5 Submit replicated mutation data

6.2.5.1 SQL Script generating replicated mutation table

A replicated mutation table was created to store mutation data that have previously been found in a population and being found in another population recorded in publications or submitted by authentic researchers as shown in Figure 21.

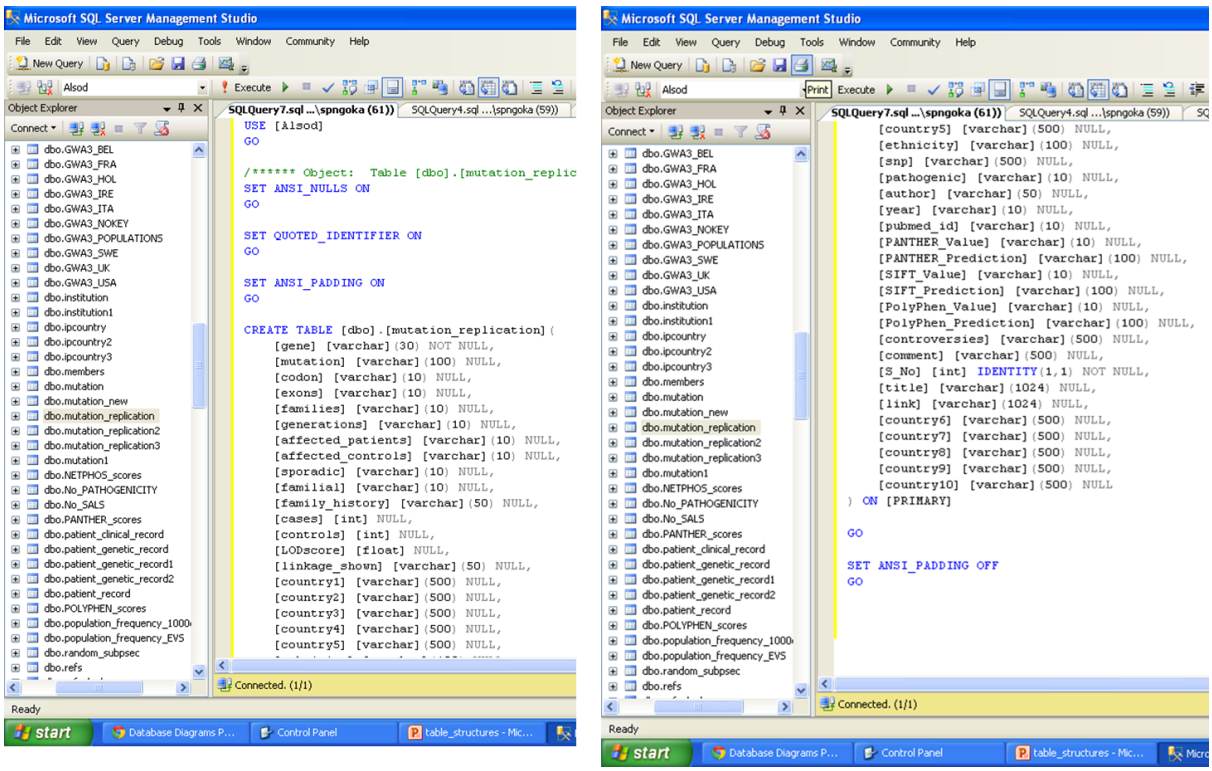


Figure 21: Replicated mutation table on SQL server

6.2.6 Submission process into ALSoD

I designed a single page in Figure 22 for entering replicated mutation data. The data curated here are gene mutations that have been submitted before but discovered in another population by another researcher and published by another author.

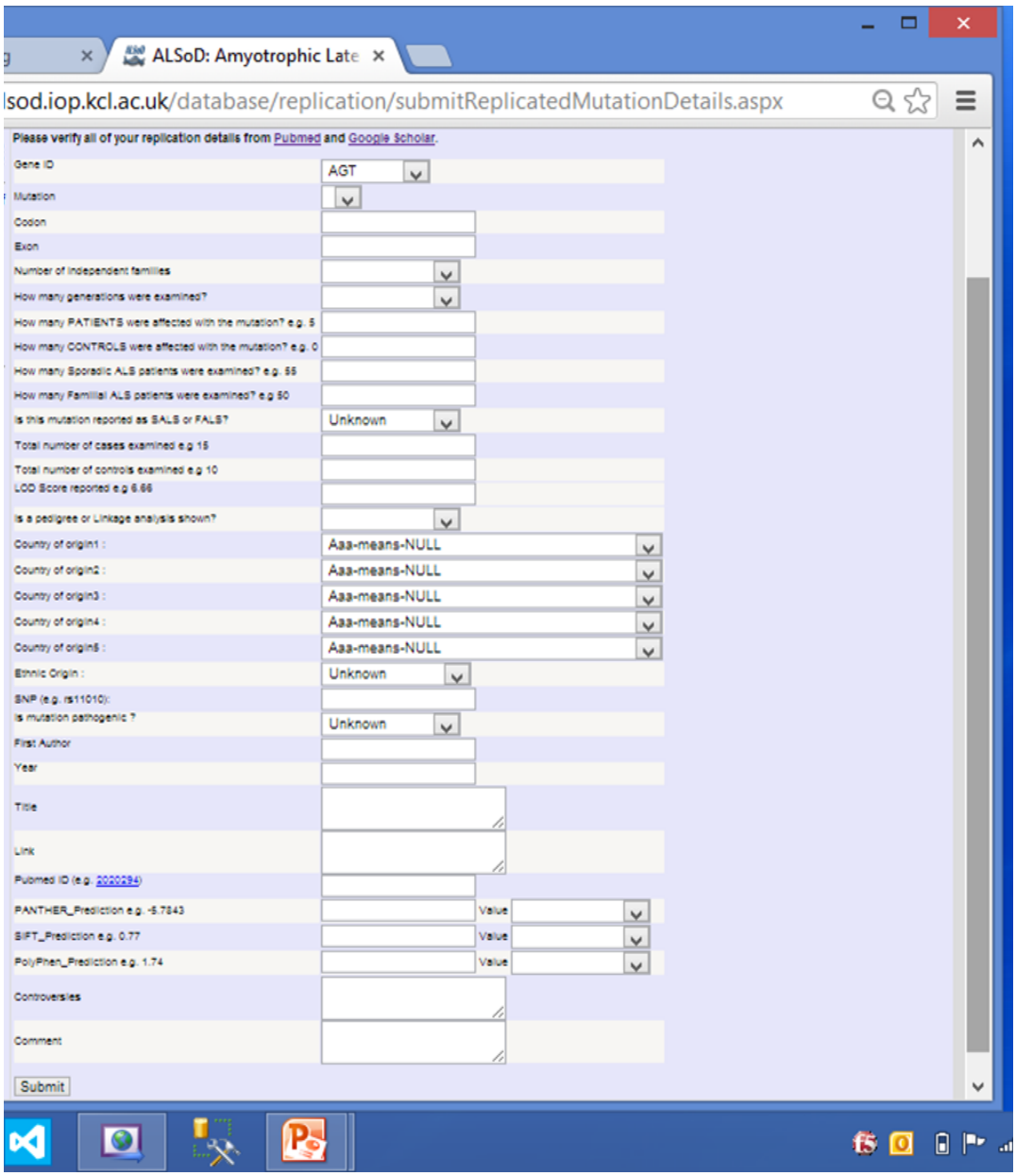


Figure 22: Replicated mutation submission page

Click on submit button and the data are automatically inserted into the ‘mutation replicated’ table of the SQL database.

6.2.6.1 Database Schema

The design of the database involves various tables in the database schema from storing data on sequence mutation data on genes to anonymized patient data including phenotypic features and country of origin. The

tables are the basic foundational blocks of the database from which queries are carried out to produce results. Here is the list of tables and how they are connected in Figure 23 with a larger view on Appendix 40.

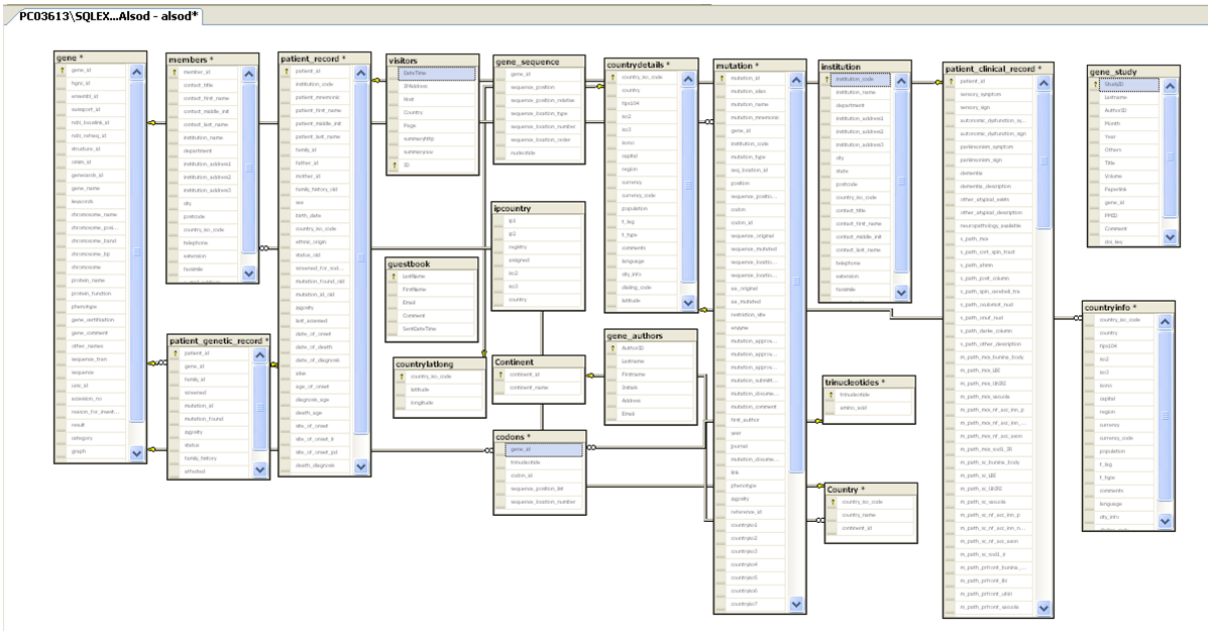


Figure 23: Database Schema

6.2.7 Search terms on search engines

Published and unpublished data are continually curated by the database administrator and registered ALS researchers. Published data are obtained by regular searches from PubMed and Google Scholar, and screening of cross-references from relevant publications. The protocol for data extraction is shown in Appendix 6 - Data extraction from publications.

6.2.7.1 Pubmed

After a long search for the most encompassing search term which would not miss out publications required and which would not include too many irrelevant publications, I was able to get these keywords together.

(mutation[Title] OR variant[Title] OR novel[Title]) AND (Amyotrophic Lateral Sclerosis[Title] OR Motor Neuron Disease[Title] OR ALS[Title]) AND 2013[All Fields]

6.2.7.2 Google Scholar

The same process used in the pubmed search above applies to the Google search and I came up with theses search terms:

novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" and select Year as ‘2013’

6.2.7.3 Automated search

Email Alerts are sent regularly from Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration journal to alsonlinedatabase@gmail.com.

News page http://alsod.iop.kcl.ac.uk/Search/searchWeb.aspx displays current research information using "New gene mutation Amyotrophic Lateral Sclerosis ALS" as the search term on the Google search plug-in attached to ALSoD.

http://alsod.iop.kcl.ac.uk/misc/Top10.aspx displays the last 10 mutation and patient data submitted.

6.2.8 Contributors and Collaborators

Unpublished data are uploaded by users (as described above) or made available to the ALS investigators by collaborations, and are published on the ALSoD website after consistency checks are performed by the database administrator.

6.2.8.1 Contributors

These are found on webpage http://alsod.iop.kcl.ac.uk/misc/contributors.aspx. Connections are made using the hyperlink.

6.2.8.2 Collaborators

Some of the tools for analysis of data were made possible through partnership with external researchers found on the webpage http://alsod.iop.kcl.ac.uk/misc/resources.aspx.

6.2.9 Unique Identifiers

To allow integration of different databases through one gateway, ALSoD uses unique identifiers to systematically link to other bioinformatics tools and comprehensive databases. How they are extracted and used will be discussed later in 6.3.10. This is used especially in the gene table where ids from other databases for a particular ALS gene are stored on ALSoD database.

6.2.10 Graphical display of phenotypic data

An overview page is available for all genes when selected from the homepage. There are 6 sections on this page displaying information on the gene, a graphical display of patient data, a tabular summary of patient data for the gene, bioinformatics links to webpages specific to the gene, animal models that have been used in

research relating to the gene and key publications that have reported mutations or replicated data on the gene.

6.2.11 Mapping variants on Google Earth

The mapping of gene variants on Google Earth allows researchers or patients to view the geographical distribution of reported gene variations associated with ALS. The same computational method allows us to map the origin of users and show which countries predominantly access ALSoD. This section is discussed in more details in section 6.3.2.3

6.2.12 Meta-analysis and on-the-fly analysis of association data

For users with their own association data, an on-the-fly analysis is available to combine the data available in ALSoD with unpublished user data that can be confidentially uploaded. The user data is formatted accordingly before upload and the result is fed back in minutes without storing users' data on the database. This section is discussed in more details in section 6.3.4

6.2.13 Instruction Manual on updating database

Due to the volume of curation required in ALSoD regularly, help is required outside the team to update the database. Helpers are not allowed to submit data directly into the database to minimise errors and inconsistencies. So, spreadsheets have been prepared to allow helpers store data intended for the database. Once the data have been verified by the Administrator, an SQL query is used to insert the data into the relevant table of the ALSoD database.

6.2.14 Updating Database

Searching through pubmed for mutations and corresponding patient data from current publications between 2011 and 2013 with search criteria provided for the available 105 genes on ALSoD

Genes to concentrate on are: ALS2, ANG, DCTN1, FUS, FIG4, DAO, VCP, LUM, SETX, C9orf72, SOD1, TARDBP, VAPB, SPG11, NEFH, OPTN, PFN1, TAF15, EWSR1, SIGMAR1, UBQLN2, SQSTM1

For example SOD1, open http://www.ncbi.nlm.nih.gov/pubmed

Type: "SOD1 ALS mutation" click on 'Humans' and '5years' on the left section

Go through the list, select the relevant ones, send to 'file' and save the file.

Check the gene overview page of the selected gene e.g. http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SOD1

If the result list contains one of the publications on the key publication section of ALSoD, ignore it and move to the next on the list until there exists a publication that is not available on the ALSoD list.

For patients, open the ALSoD_update_july_2013 spreadsheet, go to patients sheet

Fill in the relevant columns

Remember to save the file

For mutations, open the spreadsheet, go to mutations sheet

Fill in the relevant columns

Remember to save the file

6.2.14.1 Add ALS animal publications to database

open http://www.ncbi.nlm.nih.gov/pubmed

Type: “SOD1 ALS” click on ‘other animals’ on the left section

On the ‘Send to’ dropdownlist, select File

Choose CSV format, Recently Added and click on ‘Create File’ button

A pubmed.csv file is displayed in excel by default then save the file to represent the gene e.g SOD1_animalmodel. If not, open excel spreadsheet then open the saved file.

On cell L1, Type ‘Year’

On cell L2, Type ‘=RIGHT(E2,4)’ then double-click the right hand bottom corner to copy and paste the formula on all cells of the column.

On cell M1, Type ‘Author’

On cell M2, Type ‘=RIGHT(K2,LEN(K2)-SEARCH("|",K2,3))’ then double-click the right hand bottom corner to copy and paste the formula on all cells of the column.

On cell N1, Type ‘Gene’

On cell N2, Type gene name ‘SOD1’ then double-click the right hand bottom corner to copy and paste the formula on all cells of the column. If it doesn’t show everything as ‘SOD1’, click on the autofill option icon on the right hand corner of the column and select ‘Copy Cells’

Create sheets for mouse(or mice), zebrafish, chicken, chimpanzee, rat, cow, dog and drosophila (or any other you notice). Then rename them accordingly eg. Mouse

On the main SOD1_animalmodel file, take the cursor to A2, Click the Data menu and click on Filter.

Click on the dropdownlist button, Select ‘Text Filters’, select ‘contains’

Type (‘mouse’ or ‘mice’)

Select the top left hand box to highlight the entire sheet, right-click on the box and select copy or CTRL+C, open the ‘mouse sheet’ and paste values or CTRL+V.

On cell O1, Type ‘Model’

On cell O2, Type model name ‘mouse’

Repeat from (m) for each of the models until done

Go back to (a) for another gene.

6.2.14.2 Adding data from replicated studies using the available online submission forms

On the left hand side of the homepage, register to get a username and password.

For temporary access, use ‘Tester’ and ‘Testing’ for username and password respectively.

Select submit a replicated mutation link

Fill in only details of mutations available in the list and details available in publication.

Click ‘submit’ button when completed.

6.2.14.3 Completing animal model file with missing data

Open the ‘animalmodels.xls’ file

Go through each row from the beginning and fill in the missing data for columns Mouse MGI ID, HGNC ID and OMIM Gene ID.

Take note of the copy and paste using autofill option by remembering to use ‘copy cells’

6.2.14.4 Finding rsid for variation in the human genome using HGVS nomenclature

Have a list of mutations recently added to the website that do not have complete ids like the NM or NP.

Open the Human Variation website: http://www.ncbi.nlm.nih.gov/projects/SNP/tranSNP/tranSNP.cgi to search, annotate and submit sequence.

Type the NM id e.g .NM_006940.4:c.T193>C on the Input an HGVS name and follow the screen dump process in Appendix 6.

6.3 ALSoD: A user-friendly online bioinformatics tool for amyotrophic lateral sclerosis Genetics

6.3.1 Funding and Sponsorship

The ALSoD database is a joint project of some organizations like MNDA, ALSA, WFN, ENCALS. To get the logos displayed and linkable to their corresponding websites, the ImageMap function was used. Below is a screen print of the codes used to develop a clickable image on the MasterPage.

6.3.1.1 WFN-ALS

Name: World Federation of Neurology-Amyotrophic Lateral Sclerosis

Website: http://www.wfnals.org

Logo:

Figure 24: WGN-ALS logo

6.3.1.2 ENCALS

Name: European Network for the Cure of ALS

Website: http://www.euromotorproject.eu

Logo:

Figure 25: ENCALS logo

6.3.1.3 ALSA

Name: Amyotrophic Lateral Sclerosis Association

Website: http://www.alsa.org/research/funding.cfm?CFID=153&CFTOKEN=65998082

Logo:

Figure 26: ALSA logo

6.3.1.4 MNDA

Name: Motor Neuron Disease Association

Website: http://www.mndassociation.org

Logo:

Figure 27: MNDA logo

6.3.1.5 How to create a clickable image using ImageMap control in ASP.net

On ASP.net, a server side control called <asp:ImageMap> allows you to create an image where some parts of the image is clicked and links to a desired web page. For example, I made the ALS-FTD diagram shown on the home page clickable.

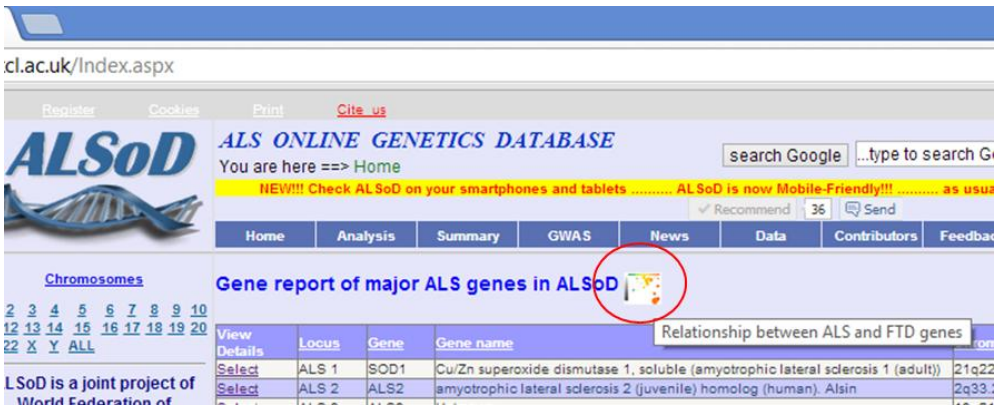


Figure 28: Clickable image using ImageMap control

To use the Paint Application on Windows 8 (or any windows OS) navigate to Start -> Programs -> Windows Accessories -> Paint.

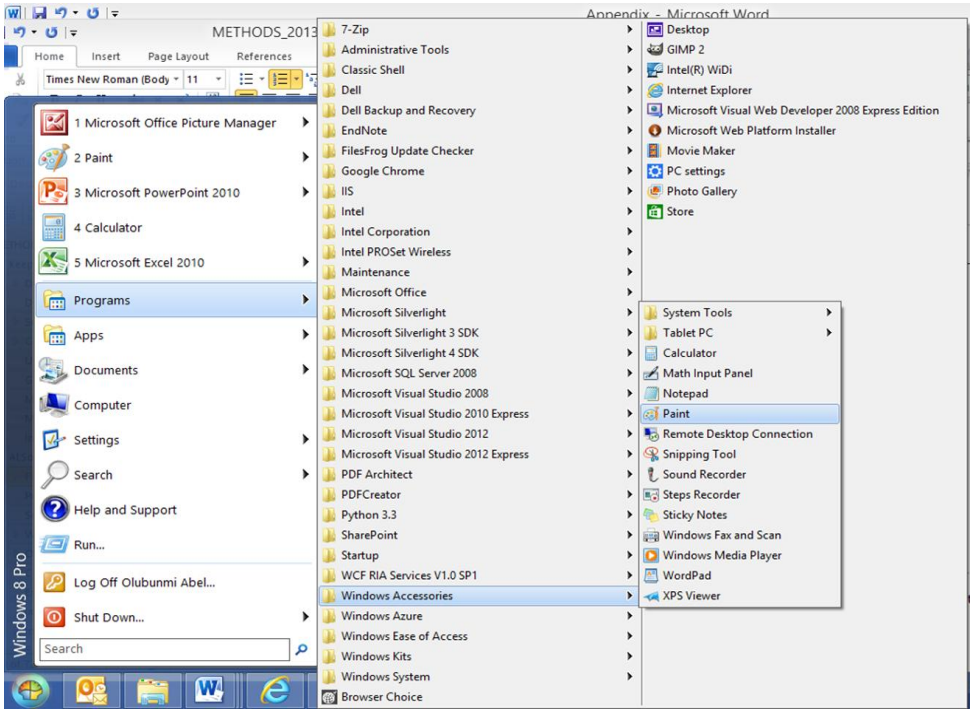


Figure 29: Opening the Paint application

There are three ways of getting hot spot regions within the image

<asp:CircleHotSpot>

<asp:RectangleHotSpot>

<asp:PolygonHotSpot>

I decided to use <asp:RectangleHotSpot> instead of the <asp:CircleHotSpot> because of the overlap of circles in the original image. Open the filename in Paint and get the hotspot coordinates for a gene like MAPT:

Left="193"

Top="331"

Right="222"

Bottom="355"

NavigateUrl=http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=MAPT

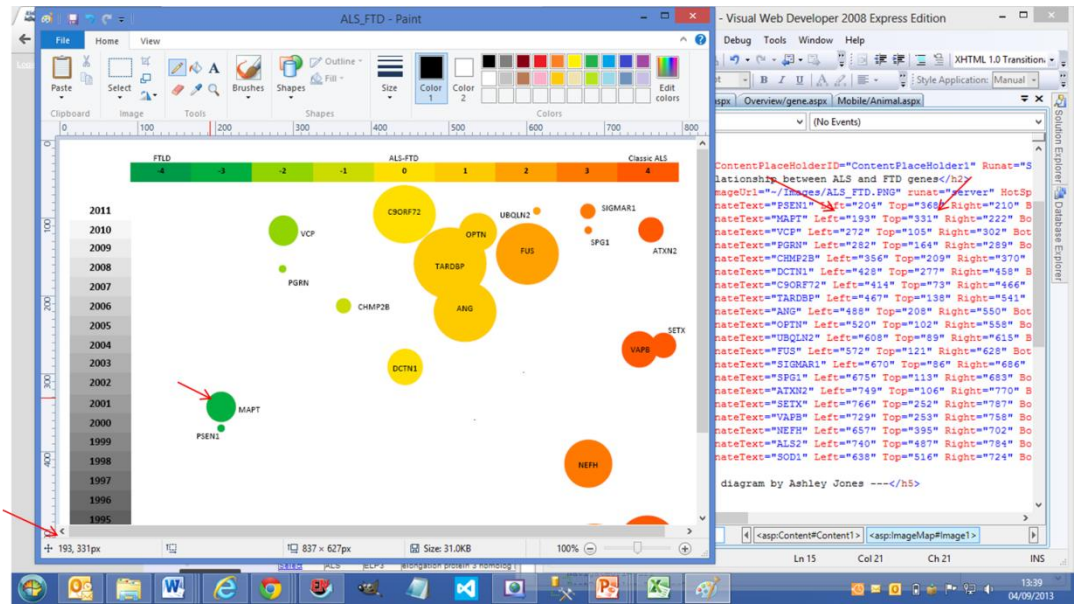


Figure 30: Using <asp:RectangleHotSpot> control

On ASP.NET, the full codes are as seen in Figure 31

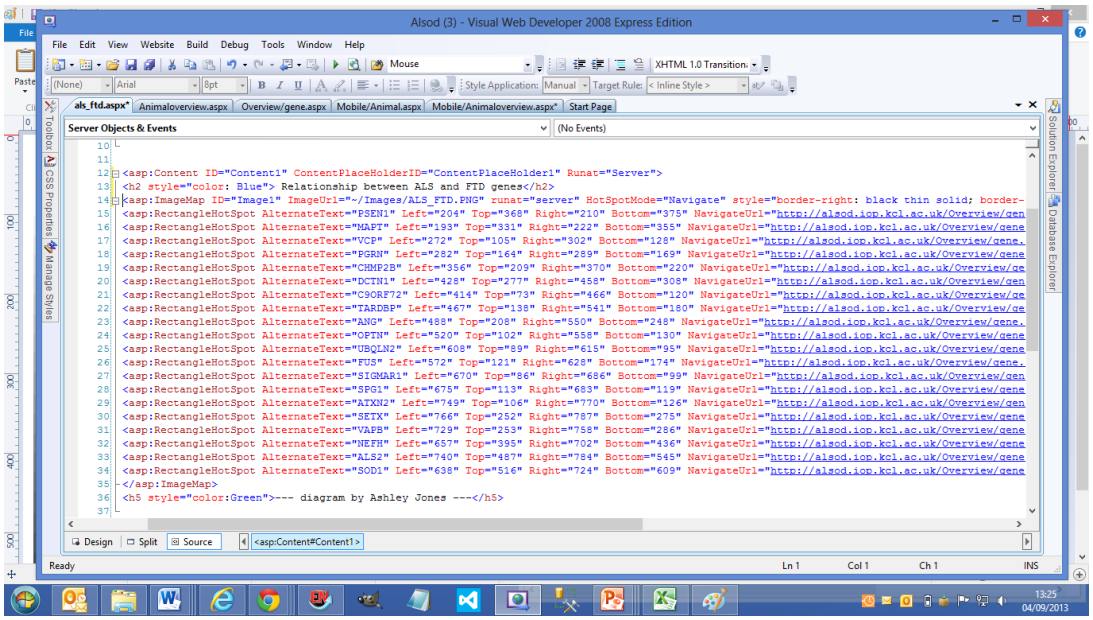


Figure 31: Full code for mapping image

Click on the gene (MAPT) and it opens up a new window displaying the gene overview of that gene in ALSoD.

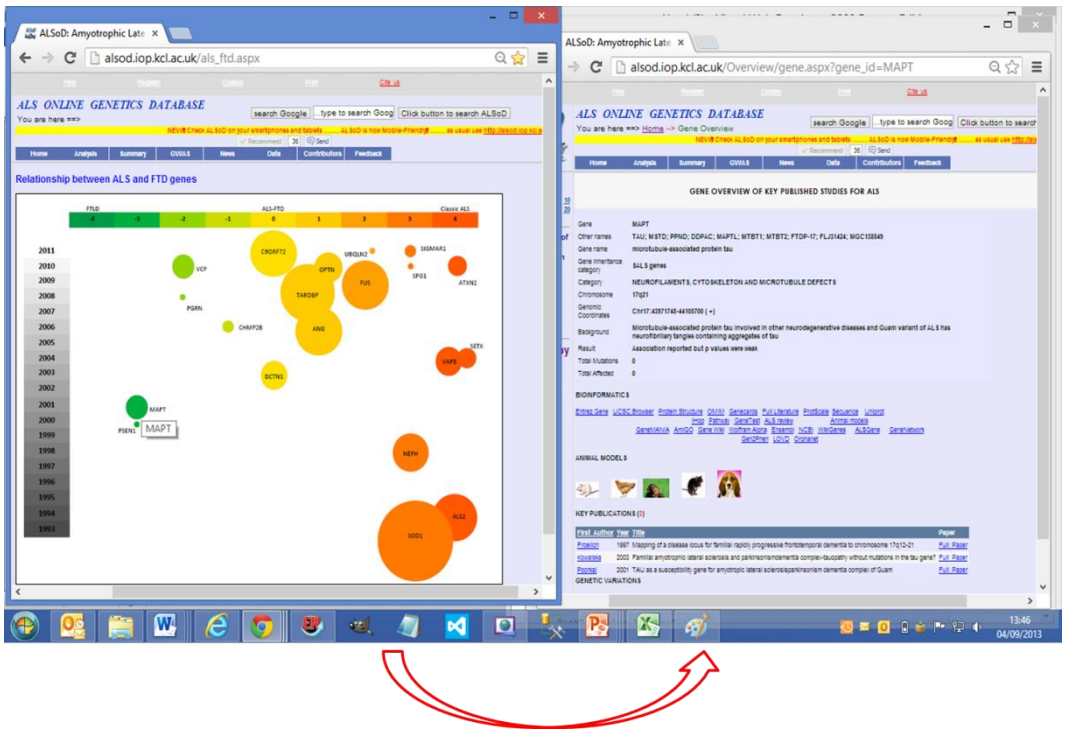


Figure 32: Clickable image showing gene overview of MAPT

The same technique was used to make the sponsor logos clickable on the MasterPage. The ImageMap section used to derive this is as below in Figure 33.

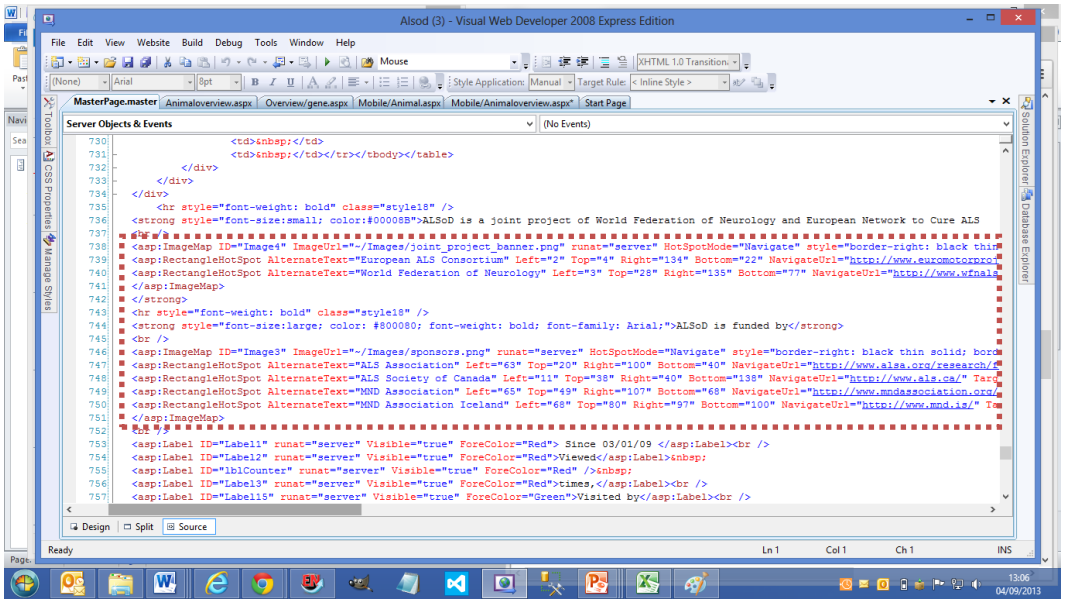


Figure 33: Using <asp:RectangleHotSpot> for logos

6.3.2 Programming

Open source programming software such as JavaScript, C#, T-SQL, Perl, XML, and VB.NET integrated under the ASP.NET platform are implemented to write codes and scripts.

6.3.2.1 Microsoft .NET framework

ALSoD uses the Microsoft .NET framework and Microsoft Visual Web Developer 2008 Express Edition to develop the user-interface dynamic Web pages.

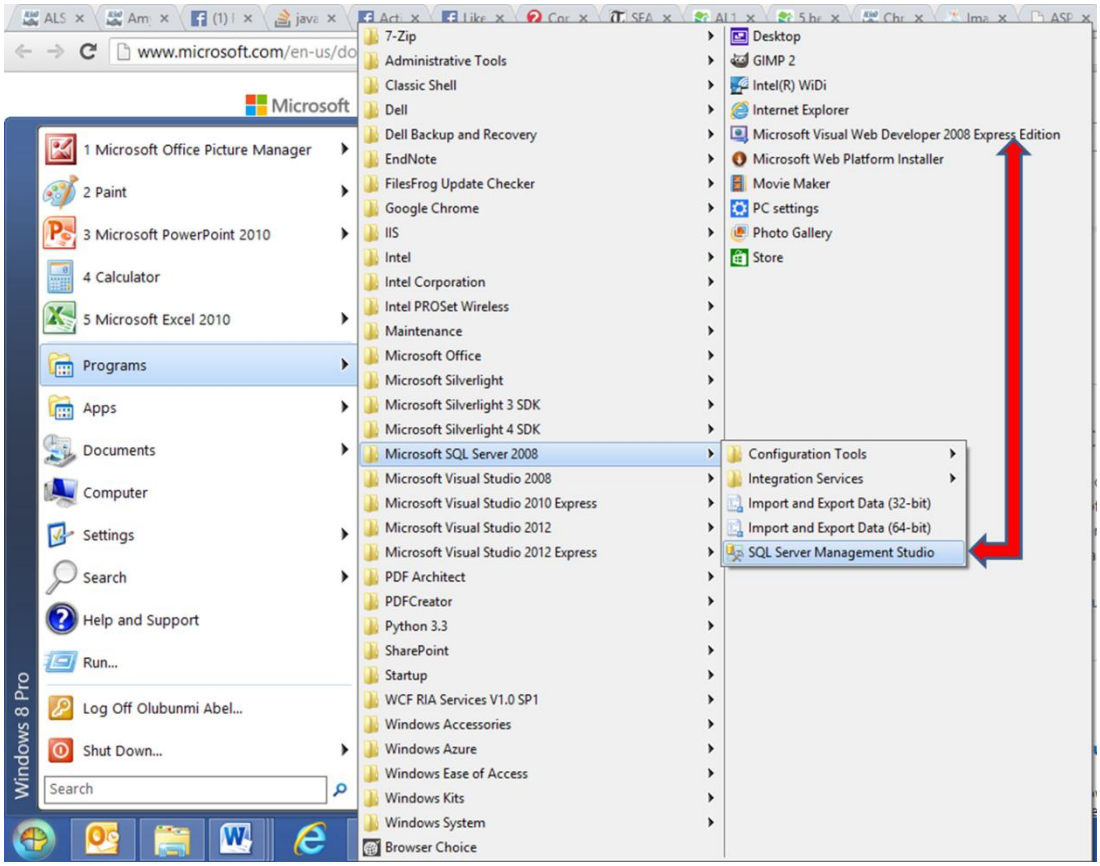


Figure 34: Calling Microsoft .NET framework

6.3.2.2 Microsoft SQL server 2008

Microsoft SQL server 2008 is used to manage the database stored on the VM3 server of the Institute of Psychiatry, King's College London but linked to the developer's PC as displayed above. Below is the description on how to link to the SQL server on VM3 provided the softwares have been installed on the developer's machine and the Institute's database administrator has configured the VM3 accessiblility using a college-recognized username and password.

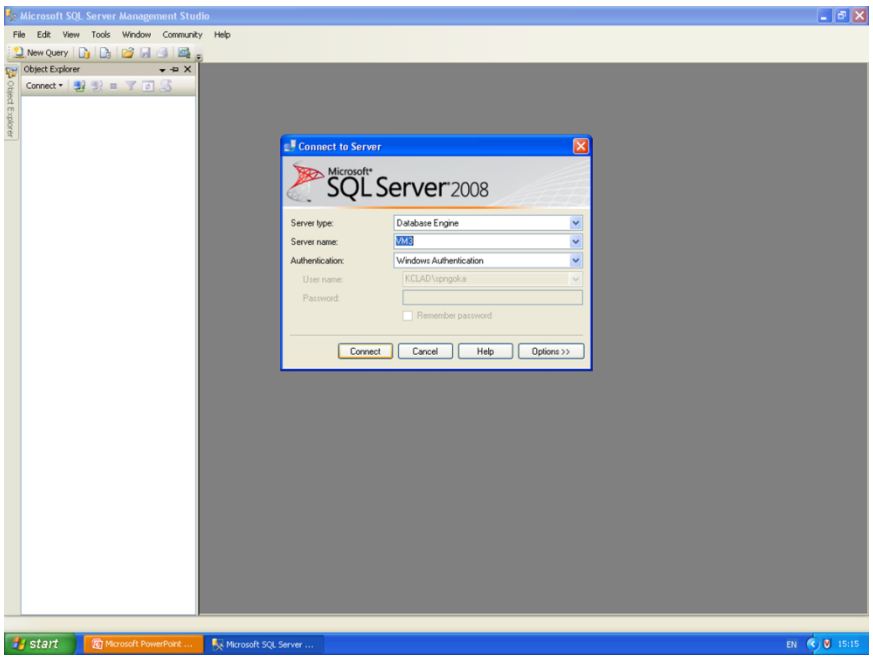


Figure 35: Microsoft SQL Server 2008 startup page

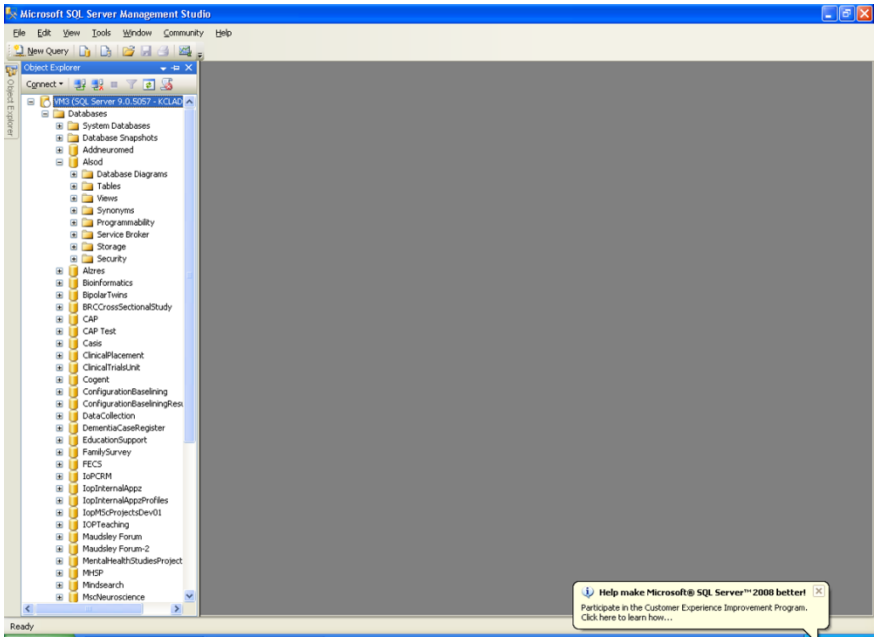


Figure 36: Microsoft SQL Server 2008 application

6.3.2.3 Google Earth API

Google AJAX search and Google Earth API have been combined to overlay geographical mutation information of ALS-related genes on the globe as seen below. Browsers used to test the application for purpose of compatibility to a large extent are Firefox, Google Chrome and Internet Explorer.

6.3.2.4 Eclipse software

The most recent programming skill I acquired through personal study is the Eclipse software which is a Plug-in allowing a mobile device programmer to develop, test and debug a Java application. Eclipse software was utilised primarily to develop an app for ALSoD. It had a very long technical process and required a lot of intricacies to finally get it working on the computer. More details on mobile application development in ALSoD are described later in section 6.5 and the whole process in form of screen dumps is in Appendix 10.

6.3.2.5 Perl

I learnt how to write scripts in Perl to manipulate large datasets. PolyPhen, SIFT and PANTHER data were restructured to suit the SQL database tables. Some codes for transposing a sequence are as shown in Appendix 22.

6.3.2.6 Python

I installed python 3.3 on my machine to run the population frequency and enable me map the mutations on ALSoD to 1000 genome and EVS. Even though the codes were written by my colleague to enable me achieve a true result, I used my programming skills to decode the script and amended it to suit the purpose. The script is in Appendix 21.

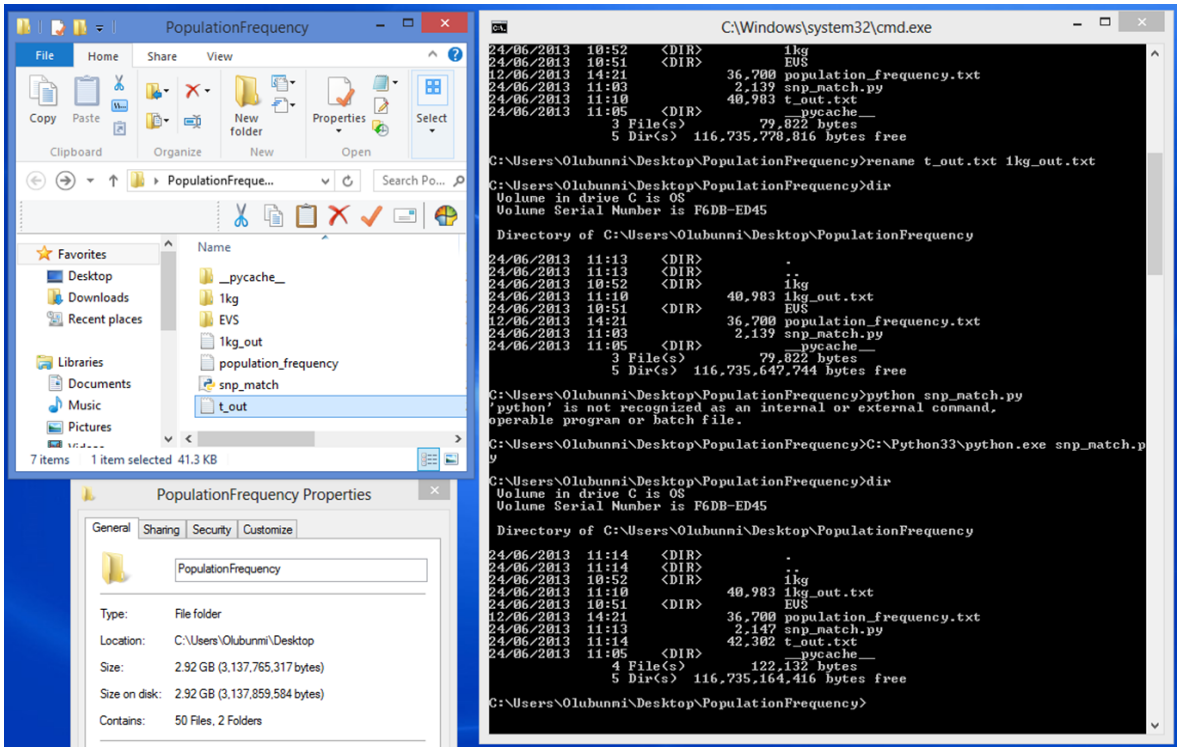


Figure 37: Using Python to map population frequency

6.3.3 Web Design

With the schema restructured, a facelift was given to the Web page and some pages were redesigned for better visual representation of data. The Graphical User Interface allows data to be interpreted and viewed at a glance instead of using the tabular format of viewing data.

6.3.3.1 ALSoD Homepage Timeline (2007-2013)

The database began in 1995 without a website but a database for storing genetic data of SOD1 mutations. In 1999, the increase in the use of technology especially the easy access to the internet propelled a collaboration work of SOD1 mutations with the ALS scientific community to develop a uniform centralized database accessible through the internet in real time. Under the web address www.alsod.org, the database was made available to the public but was short-lived due to domain name registration and web host issues. Collaboration with the Institute of Psychiatry and the sponsors of the database established the current web address http://alsod.iop.kcl.ac.uk which is still running to date.

To see how the website metamorphosed at different stages of development, ALSoD was captured by the crawler tool of an internet archive website called 'WayBackMachine', I queried the website on http://archive.org/web/web.php. Some of the unavailable views were taken from archived old files.

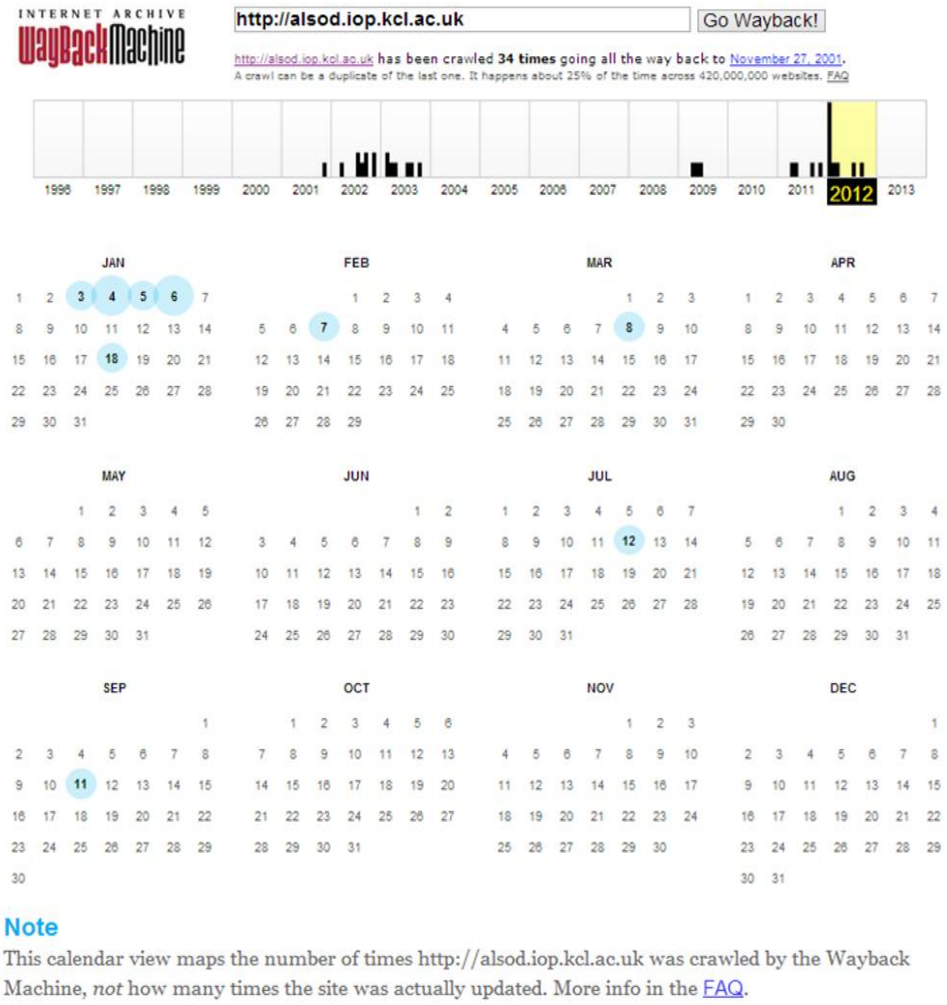


Figure 38: WayBackMachine showing ALSoD homepage timeline

6.3.3.1.1 Year 2007-2008

I took over the development and maintenance of the then 'ALSOD' database from Dr Richard Wroe in July 2008 with no working knowledge of .NET framework software or SQL Server database. I was only familiar with HTML for developing a website and Microsoft Access for databases. I was confident that it would not take too long to learn these new softwares as the basic structures are almost the same. So, for the first six months, I went through the structure of the SQL database and the website, I installed softwares on machines from scratch and I played around with some of the codes. My role was to implement the future plan for ALSOD.

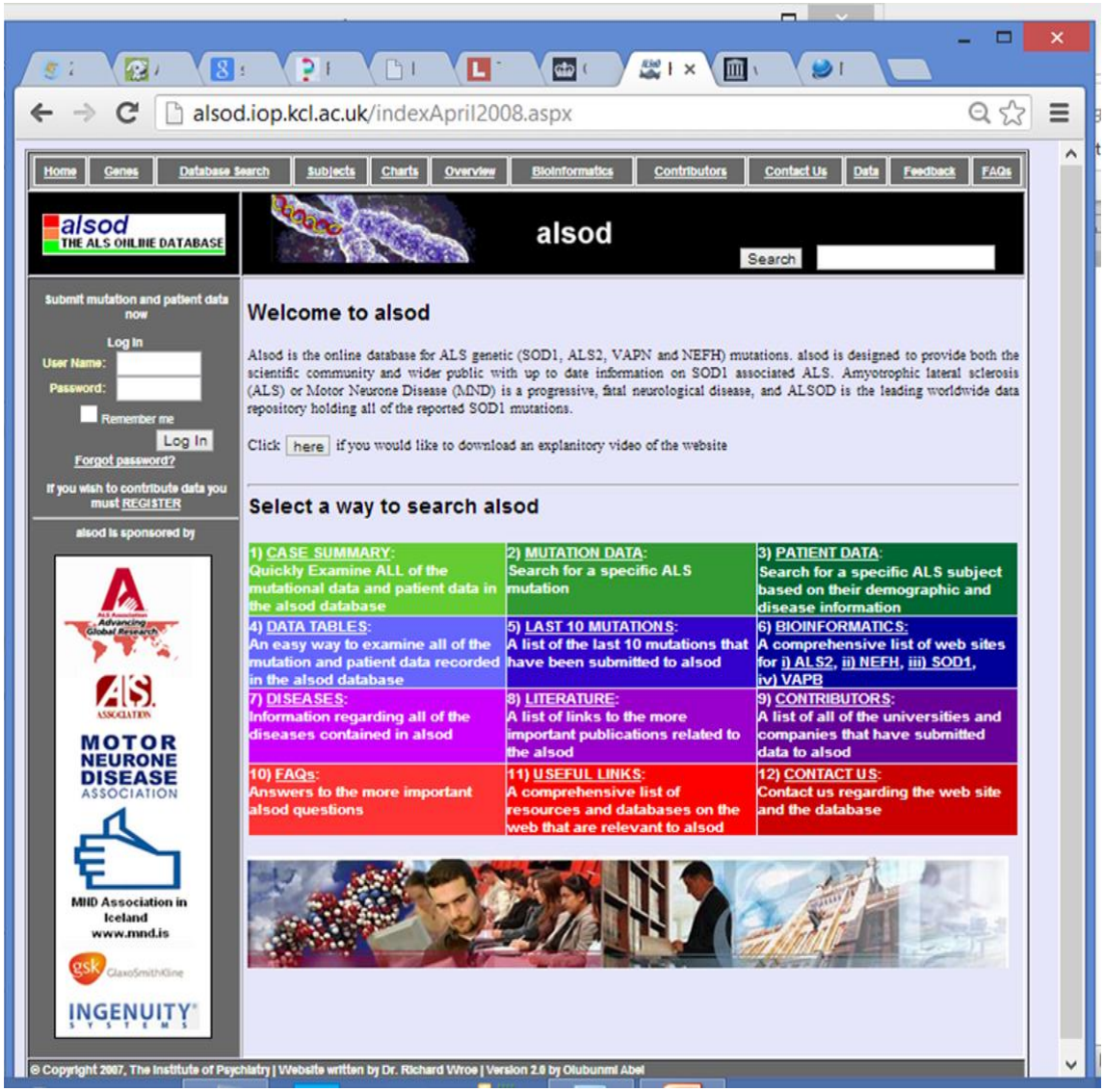


Figure 39: Home page 2007-2008

6.3.3.1.2 Year 2009

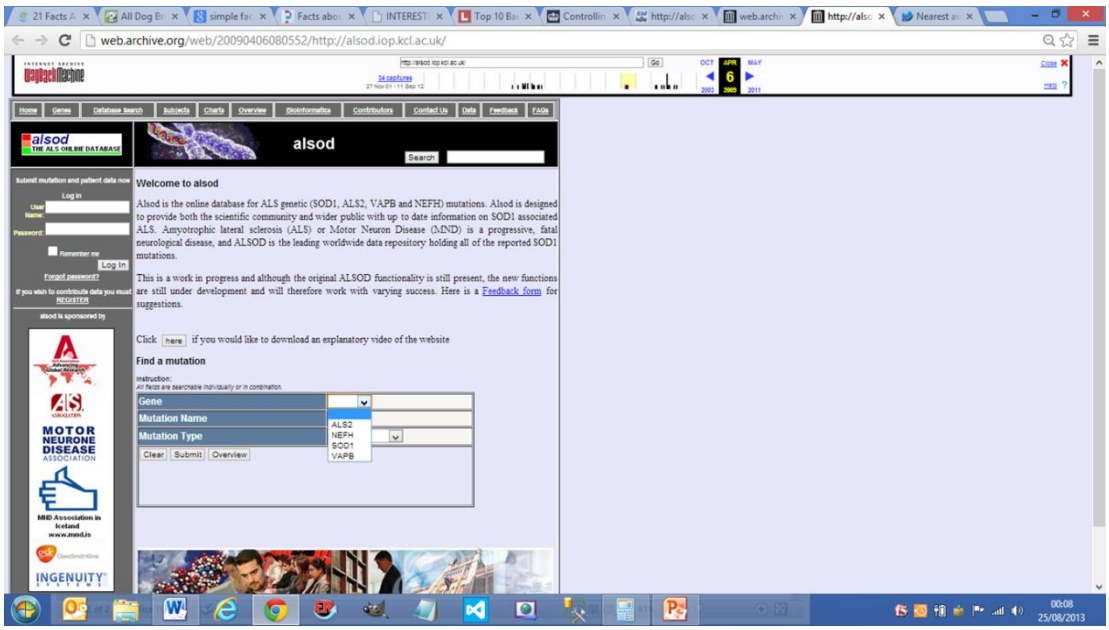


Figure 40: Home page 2009

6.3.3.1.3 Year 2010

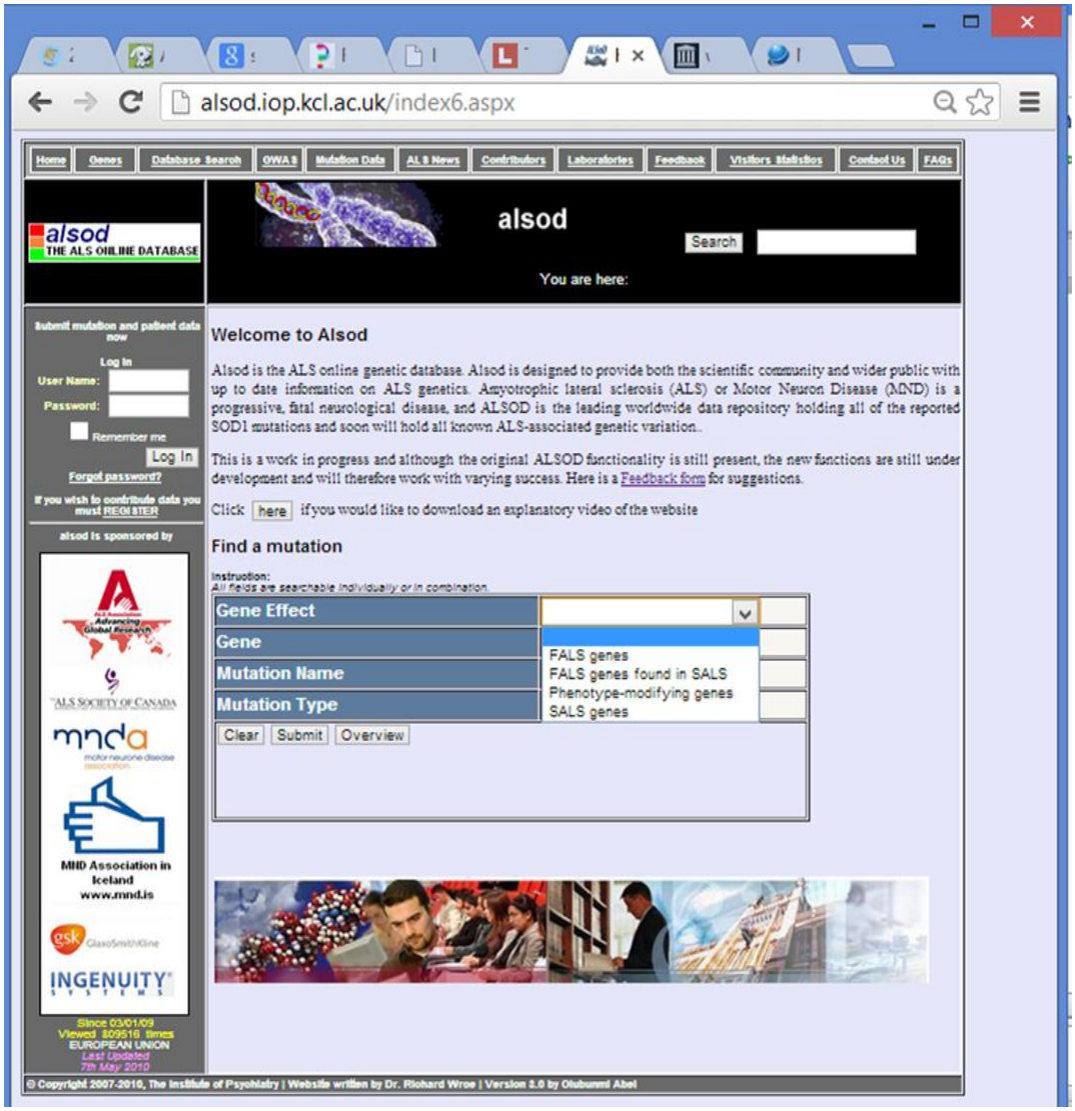


Figure 41: Home page 2010

6.3.3.1.4 Year 2011

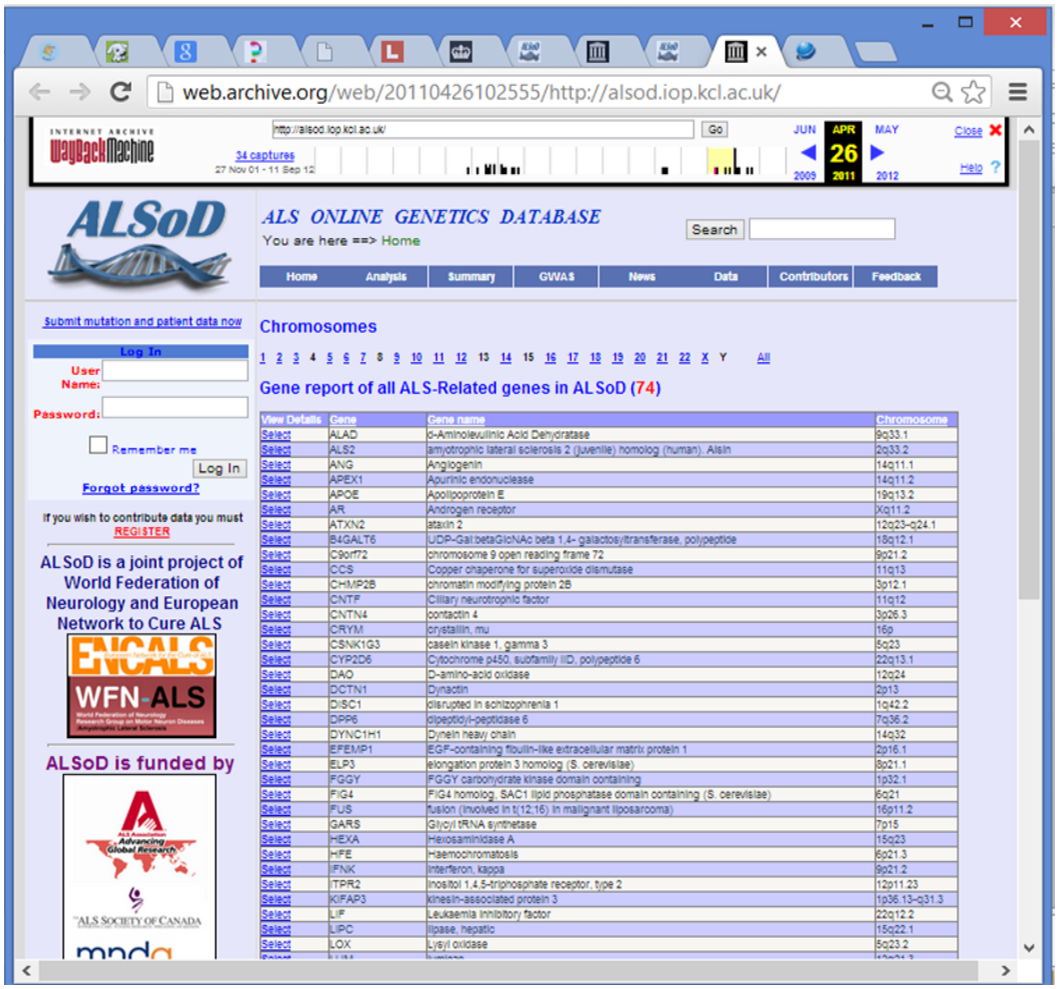


Figure 42: Home page 2011

6.3.3.1.5 Year 2012

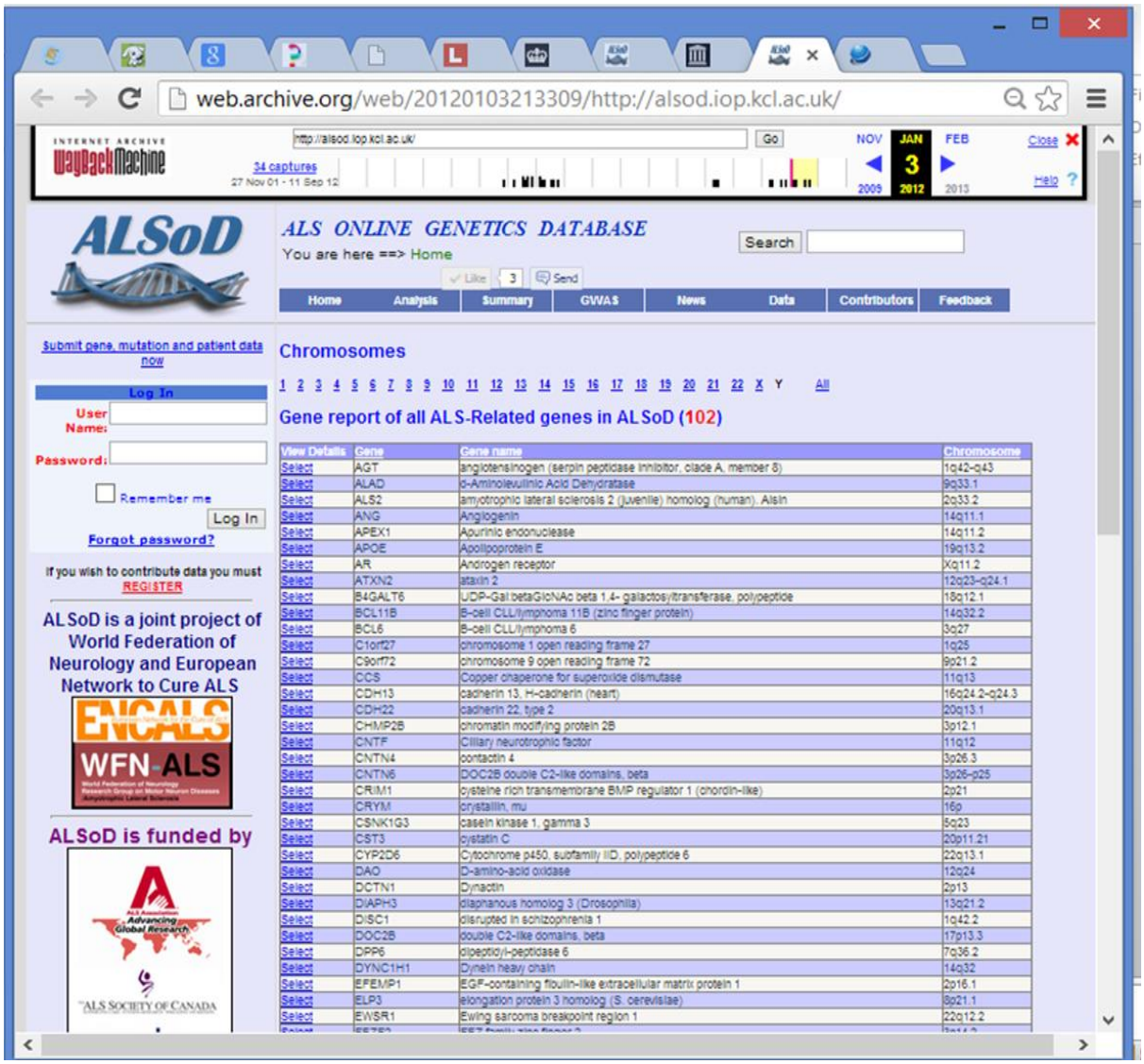


Figure 43: Home page 2012

6.3.3.1.6 Year 2013

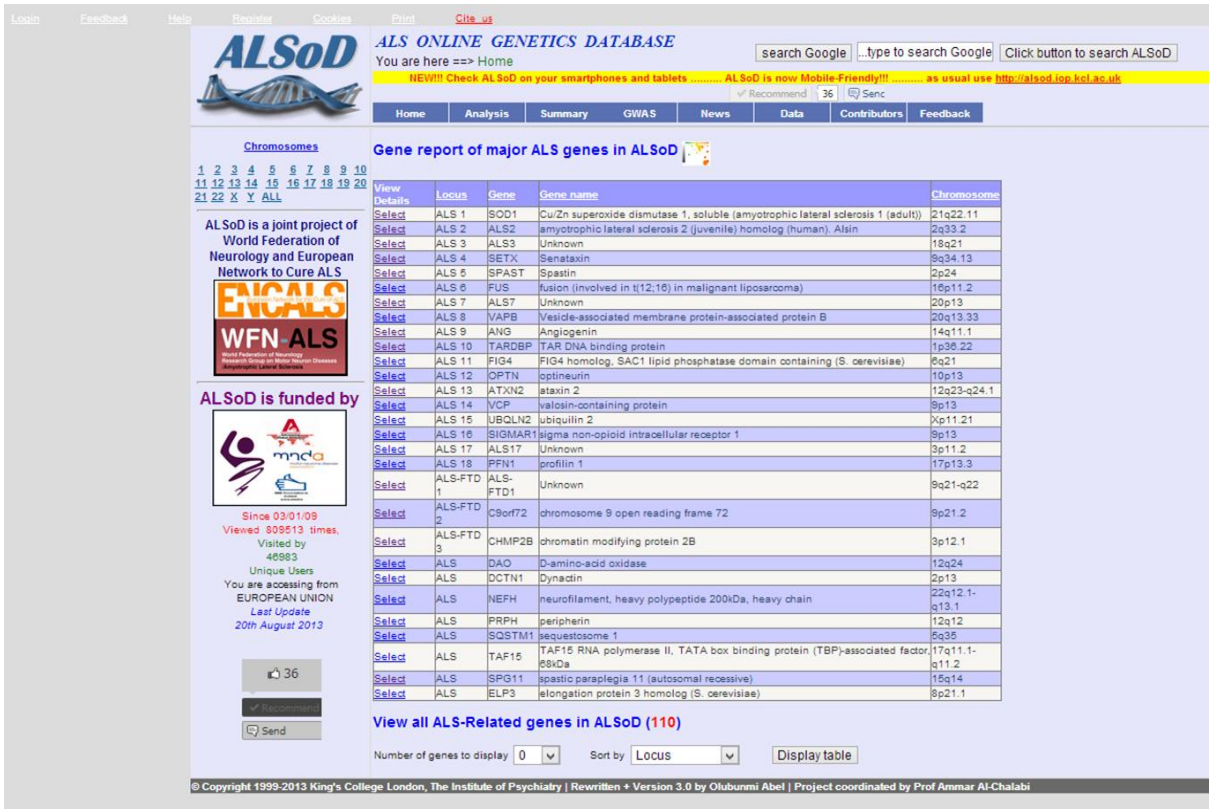


Figure 44: Home page 2013

6.3.3.2 MasterPage- A pattern for other pages with common layout and functionality.

The layout of the master page is a greyish-background table divided into five: Row1-Column1 (Header), Row2-Column1 (Top), Row3-Column1 (Left), Row3-Column2 (Contentholder) and Row4-Column1 (Footer) as seen in Figure 45.

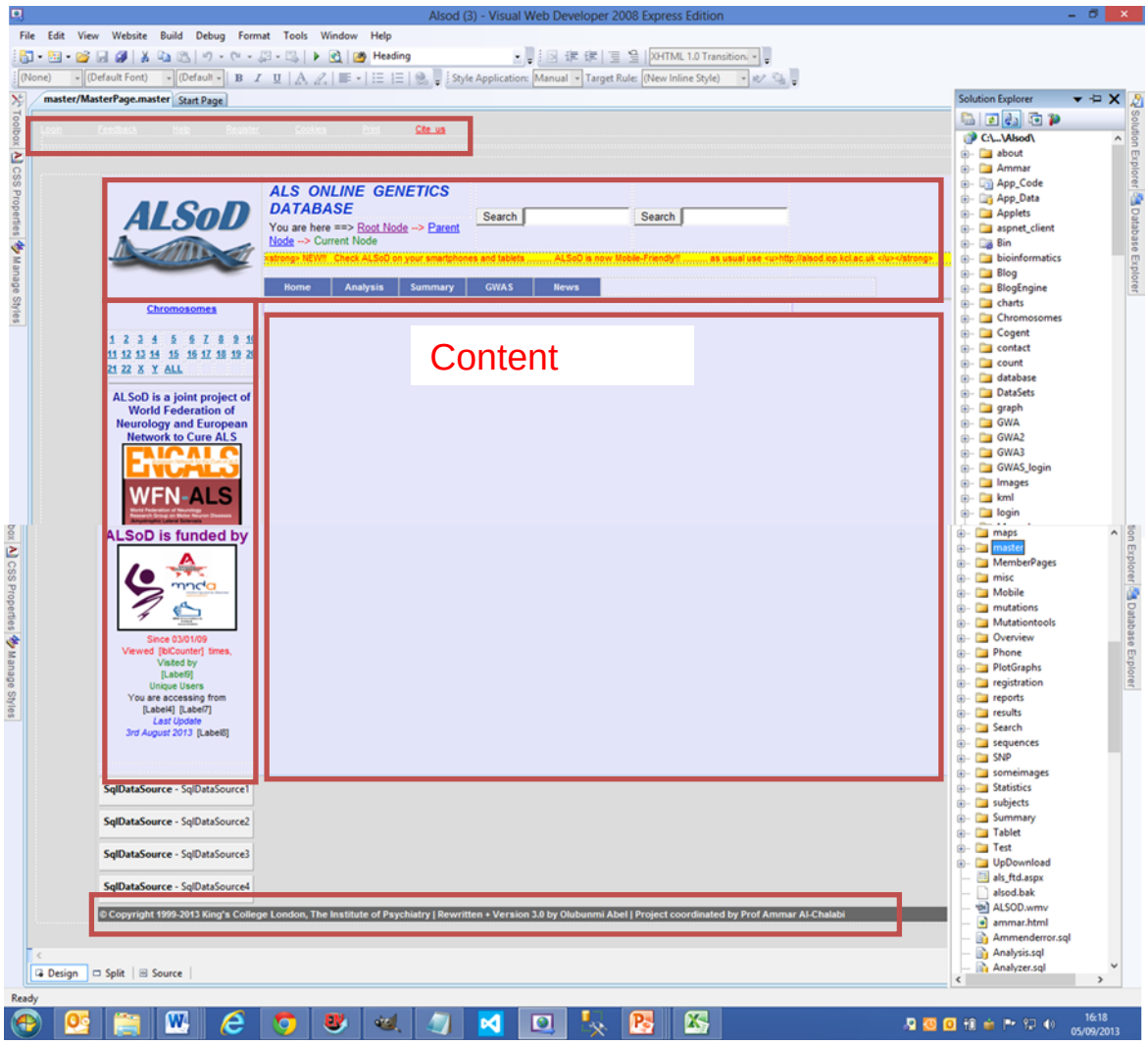


Figure 45: Masterpage structure

6.3.3.2.1 Header

Figure 46: Masterpage Header

Seven hyperlinks are displayed on the topmost part of the website for constant view and use by a user of the website.

Login - http://alsod.iop.kcl.ac.uk/login/loginAuthenticate.aspx to submit gene, mutation and patient data.

Feedback - http://alsod.iop.kcl.ac.uk/contact/guestbook.aspx to help us with feedback from users.

Help - http://alsod.iop.kcl.ac.uk/misc/FAQs.aspx to provide users with help required using our website.

Register - http://alsod.iop.kcl.ac.uk/login/preSign0.aspx to enable researchers contribute data to the ALSoD database, registration is required.

Cookies - http://www.aboutcookies.org/default.aspx?page=5 to make users of the website aware of ALSoD's cookie policy.

Print - javascript:window.print() to allow users print a page on the website by a click of a link rather than copying and pasting.

Cite us - http://alsod.iop.kcl.ac.uk/contact/contact.aspx to point user to the right direction on how to cite our work in publications.

6.3.3.2.2 Footer

Figure 47: Masterpage Footer

A statement showing the copyright, duration, host, writer and coordinator is displayed. It reads: *"© Copyright 2007-2013 King's College London, The Institute of Psychiatry | Rewritten + Version 4.0 by Olubunmi Abel | Project coordinated by Prof Ammar Al-Chalabi".*

6.3.3.2.3 Top

Figure 48: Masterpage Top section

ALSoD logo has a size of 49KB with 318 by 177 pixels designed on PowerPoint and stored as "~/Images/alsodlogo.png". It is displayed as a linkable image redirecting a user to the default or index page when clicked and reduced to 186px by width.

The name of the database "ALS ONLINE GENETICS DATABASE" is shown and a sitemap showing a breadcrumb of where a user is on the website. The sitemap is stored as a file (Web.sitemap) with details available in Appendix 13. A line of solid yellow fills highlighting a new message is shown e.g. "`Check ALSoD on your smartphones and tablets`". It also consists of the menu leading to submenus explained below.

The first search button allows a user easy access to search through the NCBI (National Center for Biotechnology Information) website for information on search term typed in the textbox.

```
protected void SearchSite(object sender, EventArgs e)
   {
      while (q.Text != "")
      {

Response.Redirect("~/Search/SearchResultPage.aspx?cx=007276566013418379788%3Ak6lhzyopur4&cof=FORID%3A9&ie=UTF-8&q=" + q.Text + "&sa=Search");

      }
 }
```

The second search button allows a user to search the ALSoD database for specific information using the Google search plug-in.

```
protected void SearchSite2(object sender, EventArgs e)
   {
      while (q2.Text != "")
      {

Response.Redirect("~/Search/SearchResultPage2.aspx?cx=007276566013418379788%3Ak6lhzyopur4&cof=FORID%3A9&ie=UTF-8&q=" + q.Text + "&sa=Search");

      }
   }
```

MENU BAR

A layout of available functions and tools can be accessed through the navigation bar and are categorized into menus and submenus below. Placing the mouse over each element displays more information about each menu as a tooltip and a dropdown submenu for more options.

Home - http://alsod.iop.kcl.ac.uk/Index.aspx

Analysis

Side-by-side comparison - http://alsod.iop.kcl.ac.uk/Statistics/statistics.aspx

Interactions - http://alsod.iop.kcl.ac.uk/overview/interaction.aspx

Summary report - http://alsod.iop.kcl.ac.uk/Statistics/report.aspx

Detailed analysis - http://alsod.iop.kcl.ac.uk/Statistics/analysis.aspx

Summary

Search - http://alsod.iop.kcl.ac.uk/Summary/summary.aspx

Summary - http://alsod.iop.kcl.ac.uk/index7.aspx

GWAS

GWAS - http://alsod.iop.kcl.ac.uk/GWA2/index.aspx

Help! - http://alsod.iop.kcl.ac.uk/misc/analysiserror.aspx

News - http://alsod.iop.kcl.ac.uk/Search/searchWeb.aspx

Data

Mutation - http://alsod.iop.kcl.ac.uk/misc/dataDownload.aspx#C1

Patient- http://alsod.iop.kcl.ac.uk/misc/dataDownload.aspx#C2

Contributors

Laboratories - http://alsod.iop.kcl.ac.uk/misc/labs.aspx

Contributors - http://alsod.iop.kcl.ac.uk/misc/contributors.aspx

Feedback

Contact/Cite us - http://alsod.iop.kcl.ac.uk/contact/contact.aspx

Visitors Statistics - http://alsod.iop.kcl.ac.uk/charts/index.aspx

FAQs - http://alsod.iop.kcl.ac.uk/misc/FAQs.aspx

Feedback - http://alsod.iop.kcl.ac.uk/contact/guestbook.aspx

6.3.3.2.4 Left

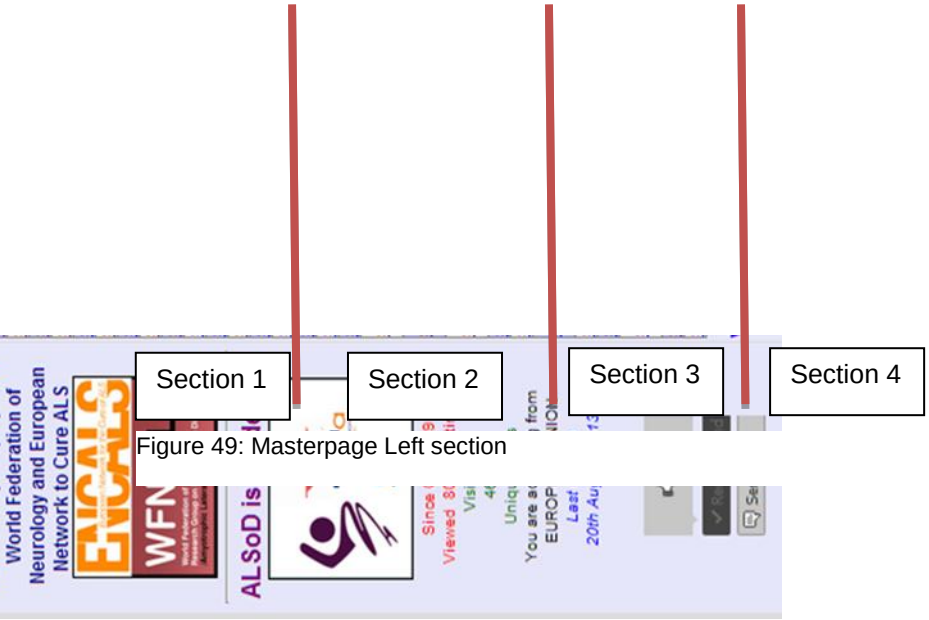


Figure 49: Masterpage Left section

The first section here is the chromosomal overview of all the genes available on the database. For example, to check for all the ALS-related genes in chromosome 9, click on 9 and a list of genes (C9orf72, VCP, SUSD1, ALAD, SETX) with arrows pointing to the location on chromosome 9 are displayed. Click on any of the linked genes and an overview page showing information for the gene is displayed.

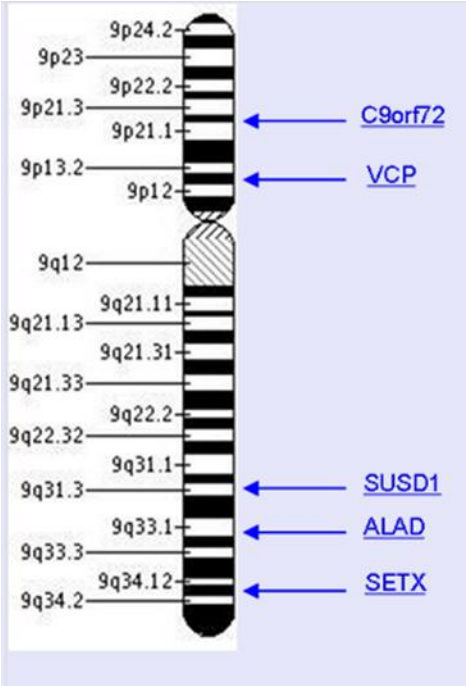


Figure 50: Chromosomal overview of chr 9

The image of the 23 chromosomes were searched and chosen online. In powerpoint application, the images were inserted on the worksheet with arrows pointing to corresponding locations on the chromosome. This indicates where the ALS gene is situated and each chromosome with the image, arrows and gene symbol are grouped and stored as an image bearing the name of the chromosome for the purpose of organising the files properly.

Imagemap function as described previously is used to link the gene symbols to the overview page of that particular gene.

```
<asp:ImageMap ID="Image1" ImageUrl="~/Chromosomes/chromo9genes.png" runat="server" HotSpotMode="Navigate">
        <asp:RectangleHotSpot AlternateText="C9orf72" Left="285" Top="84" Right="371" Bottom="96" NavigateUrl="http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=C9orf72"/>
        <asp:RectangleHotSpot AlternateText="VCP" Left="289" Top="130" Right="326" Bottom="145" NavigateUrl="http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=VCP"/>
        <asp:RectangleHotSpot AlternateText="SUSD1" Left="288" Top="379" Right="345" Bottom="392" NavigateUrl="http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SUSD1"/>
        <asp:RectangleHotSpot AlternateText="ALAD" Left="289" Top="416" Right="333" Bottom="427" NavigateUrl="http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=ALAD"/>
        <asp:RectangleHotSpot AlternateText="SETX" Left="289" Top="458" Right="336" Bottom="475" NavigateUrl="http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SETX"/>
 </asp:ImageMap>
```

The second section showing “ALSoD is a joint project of World Federation of Neurology and European Network to Cure ALS”. The work leading to these results has received funding from the European Community’s Health Seventh Framework Programme FP7/2007-2013 under grant agreement number 259867.” and the hyperlinked logos of WFN and ENCALS are visible on the website.

The third section displays an image of the funders of ALSoD which are ALSA and MNDA.

The final section shows a list of labels which are the number of times the website has been visited, the number of unique users since the monitoring code was developed in 2009, where the user of the website is accessing the website from and the last time the website was updated as seen in the flowchart in Figure 51.

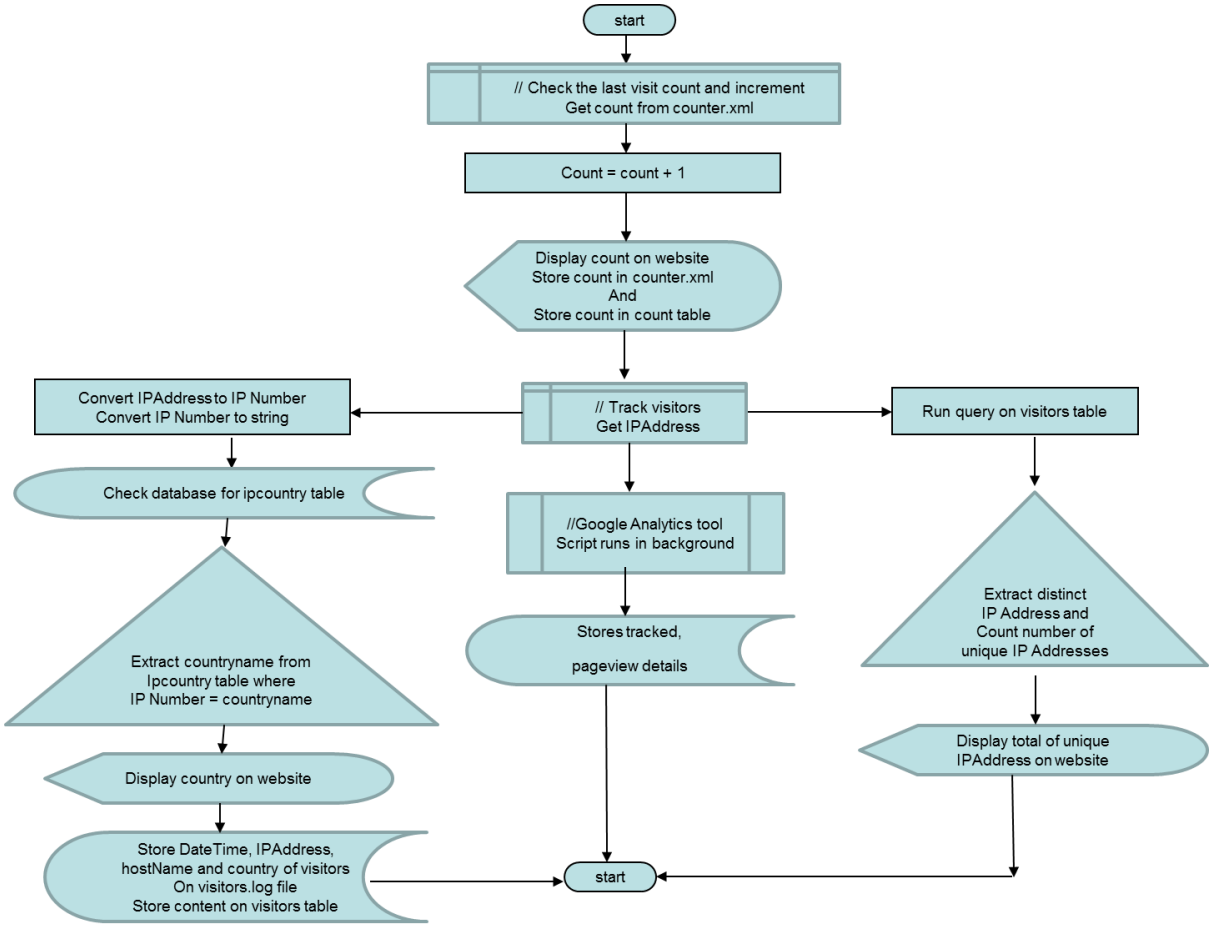


Figure 51: Flowchart for monitoring visitors

It appears on the website in form of labels whose texts are derived from a database query in real time.

6.3.3.2.5 Contentholder

This is the ContentPlaceHolder section where the content of a webpage is displayed within the content tag:

```
<asp:Content ID="Content1" ContentPlaceHolderID="ContentPlaceHolder1" Runat="Server">

</asp:Content>
```

6.3.3.2.6 Hidden codes (behind the scene)

These are codes in form of scripts for example, tracking IP addresses, search tools, mobile device detection and social media widgets on masterpage described beow:

6.3.3.2.6.1 Mobile detection

```
void CheckBrowserCaps()
  {
     //String labelText = "";
     System.Web.HttpBrowserCapabilities myBrowserCaps = Request.Browser;
     if (((System.Web.Configuration.HttpCapabilitiesBase)myBrowserCaps).IsMobileDevice)
     {
        //labelText = "Browser is a mobile device.";
        Response.Redirect("~/Mobile/index.aspx");
     }
     else
     {
        //labelText = "Browser is not a mobile device.";
        Response.Redirect("index1.aspx");
     }

     //Label1.Text = labelText;
  }

void DetectUserAgent()
{
string strUserAgent = Request.UserAgent.ToString().ToLower();
if (strUserAgent != null)
{
```

```
        if (Request.Browser.IsMobileDevice == true || strUserAgent.Contains("iphone") ||
        strUserAgent.Contains("blackberry") || strUserAgent.Contains("mobile") ||
                strUserAgent.Contains("windows ce") || strUserAgent.Contains("opera mini") ||
                strUserAgent.Contains("palm") || strUserAgent.Contains("android")
                || strUserAgent.Contains("samsung") || strUserAgent.Contains("nokia"))
        {
                Response.Redirect("~/Mobile/index.aspx");
        }
}
}
```

6.3.3.2.6.2 Track visitors

Appendix 14 displays the script for tracking users of the website by recording the IP address of the user on the database and analysing the location of the user.

6.3.3.2.6.3 Search

To search for specific ALS terms in Google Search application

```
protected void SearchSite(object sender, EventArgs e)
   {
      while (q.Text != "")
      {

Response.Redirect("~/Search/SearchResultPage.aspx?cx=007276566013418379788%3Ak6lhzyopur4&cof=FORID%3A9&ie=UTF-8&q=" + q.Text + "&sa=Search");
      }
 }
```

To search for specific keywords in ALSoD database using Google Search plugin

```
protected void SearchSite2(object sender, EventArgs e)
   {
      while (q2.Text != "")
      {

Response.Redirect("~/Search/SearchResultPage2.aspx?cx=007276566013418379788%3Ak6lhzyopur4&cof=FORID%3A9&ie=UTF-8&q=" + q.Text + "&sa=Search");
      }
```

```
}
```

6.3.3.2.6.4 Social Media

Using Like button of facebook as one of the ways of monitoring the relevance of the database to users, It allows a form of feedback and approval by users who recommend the website to a 'friend' connected to their facebook profile. It also helps to create awareness of this database to the world.

```
<div id="fb-root"></div>
<script>    (function(d, s, id) {
        var js, fjs = d.getElementsByTagName(s)[0];
        if (d.getElementById(id)) return;
        js = d.createElement(s); js.id = id;
        js.src = "//connect.facebook.net/en_GB/all.js#xfbml=1";
        fjs.parentNode.insertBefore(js, fjs);
    } (document, 'script', 'facebook-jssdk'));
</script>
```

6.3.3.3 Homepage – An overview of all genes in ALS with chromosomal view

1.1) This section shows a list of all the available genes discovered to date arranged alphabetically with the total number of the genes also displayed. There are 5 sortable columns: View Details, Locus, Gene, Gene name and Chromosome (as shown above in Figure 44)

1.3) To view the available information on a gene like FUS, scroll down to locate the FUS gene name, click on SELECT and it navigates to the gene overview page of key published studies for ALS.

1.4) Five sections displayed particularly for the gene are Brief information, graphical representation of genetic data, bioinformatics links, key publications and genetic variations.

1.5) Brief information shown are the Gene symbol, Other names, Gene name, Gene inheritance category (FALS, SALS, FALS found in SALS or Phenotype-modifying genes), Category, Chromosome, Background, Result, Total Mutations, Total affected patients recorded on the database.

1.6) Graphical representation showing a box plot on Age of Onset and 3 pie charts on Site of Onset (Limb/Bulbar ratio), Gender proportion (Male/Female ratio) and Inheritance type (FALS/SALS ratio). A single row tabular summary of the graphical representation is also displayed.

1.7) Bioinformatics links to freely available scientific websites (HGNC [525], Entrez Gene [526], UCSC Browser [527], Protein Structure[528], OMIM [529, 530], Genecards[531], ProtScale [532], KEGG [533], Uniprot [534], iHop [535], Pathway in KEGG [533], GeneTest [536], AmiGO [537], Ensembl [538], NCBI [480] , Life Science DB (Japan) [485], ALSGene [539]) and non-scientific knowledge-based websites (GeneWiki [540], WolframAlpha [541], Free online access to the Wolfram|Alpha computational knowledge engine Free online access to the Wolfram|Alpha computational knowledge engine Free online access to the Wolfram|Alpha computational knowledge engineWikiGenes [542]).

1.8) Key publications are displayed with 4 sortable columns showing the total number. The linkable first author connected to the Pubmed abstract on the NCBI website, the year of publication, title of the publication and a link to the full paper (where available).

1.9) Animal models section displays images as hyperlinks to details of animals that have been used for research in ALS. It links to information about chromosomal position, genomic coordinates, biological name, key publications with reference to pubmed and links to other animal model bioinformatics databases like Entrez Gene, UCSC Browser, Find a mice, OMIM and KEGG. The MGI id is used mainly as a key to integrate information from these databases into ALSoD.

1.10) Genetic variation section displays a 7-column sortable table on the Single Nucleotide Polymorphism(SNP id), basepair position, pvalue, first author, year, term and paper title.

6.3.3.4 Analysis webpage – Statistical analysis of mutation and patient data submitted to ALSoD from published and unpublished sources

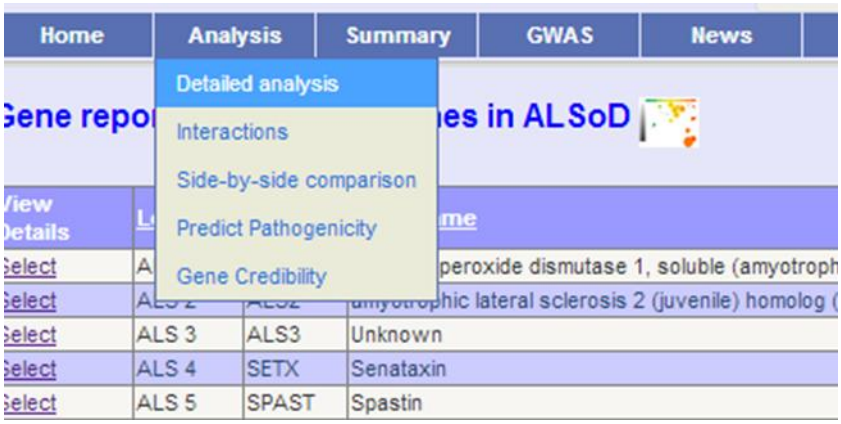


Figure 52: Analysis menu

2.1.) There are currently five web pages under the analysis menu: Detailed analysis, Interactions, Side-by-side comparison, Predict pathogenicity and Gene credibility.

2.2.) Clicking on the Detailed analysis submenu, an automatically populated checkbox list of genes with patient data, two dropdownlists on age of onset range (in years) from 0 to 100 and radiobuttons for further selections on gender, inheritance type and site of onset.

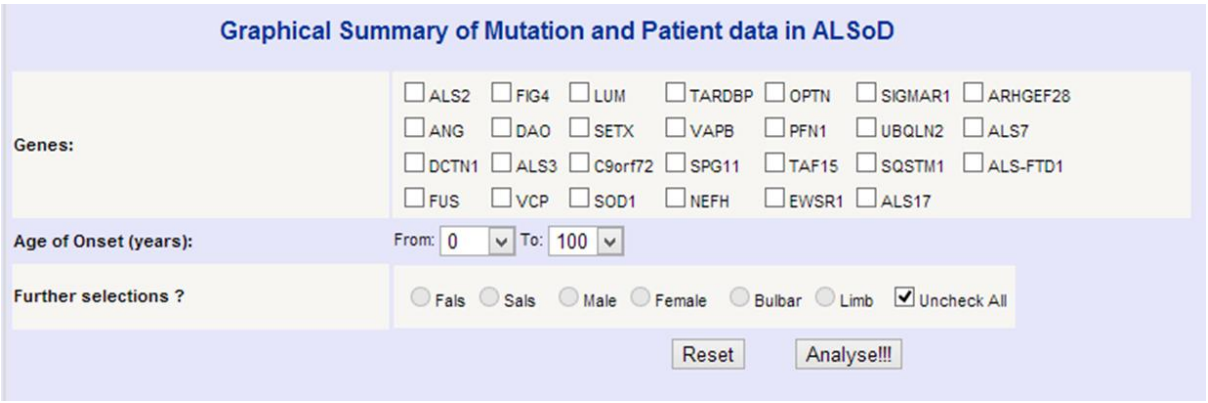


Figure 53: Detailed Analysis submenu

For example, a user wants to find out a summary detail on 3 genes TARDBP, FUS and ANG for patients with age of onset between 40 and 60. Check the boxes of the 3 genes, select 40 from the first dropdownlist and 60 from the second dropdownlist. Leave the default checkbox checked to view all selections. Once all the selections required have been done, the query is executed by clicking on the ANALYSE button. Pie charts of Gender ratio, Family ratio and Site of Onset ratio; a column chart of the mean age of onset for the 3 selected genes; a box plot of the mean age of onset, a doughnut chart of the country ratio where the affected patients

come from on the globe; and 2 tables showing specific and cumulative figures on the selected genes. Where a mistake has been made or a new set of selections is required, the RESET button needs to be clicked.

2.3.) Clicking on the Interactions submenu, a checkboxlist of all available genes in ALS are displayed. A group of checkboxes dynamically derived from the gene table is displayed to allow users select as many gene interactions they would like to analyse. Once the genes are selected, the list is presented as confirmation. The “Execute” link is clicked to display the prediction on the same page. If the “Reset” link is clicked, all the selected boxes are automatically deselected and the graphical analysis is cleared to allow for fresh analysis. GeneMANIA is a free public resource that offers a simple, intuitive web interface that shows the relationships between genes in a list and analyzes and extends the list to include other related genes. It currently Indexes 1,256 association networks containing 357,605,768 interactions mapped to 134,871 genes from 6 organisms. Its flexibility and user-friendliness allows users to interact with the web interface through ALSoD stating weights as the predictive value [509].

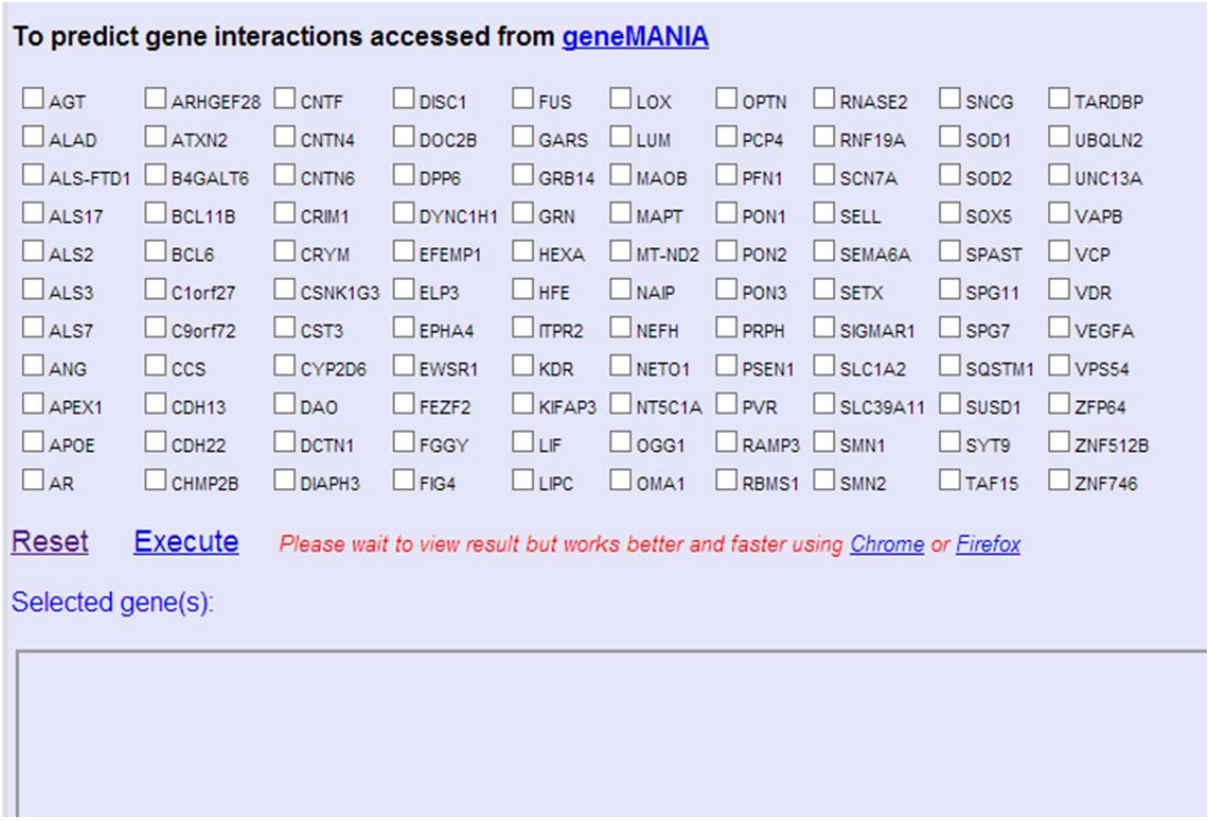


Figure 54: Interaction submenu

2.4.) Clicking on the Side-by-side submenu, runs a query analysis on genetic data displaying 4 dropdownlists to select any two genes and the age range of patients. A side-by-side comparison of mutation and patient data

on two selected genes is submitted to the database. The query is analysed and displayed graphically on the website automatically in a split of a second. It shows the site of onset, mean age using box plot, gender, family history, nationalities of patients and other joint analysis of both genes.

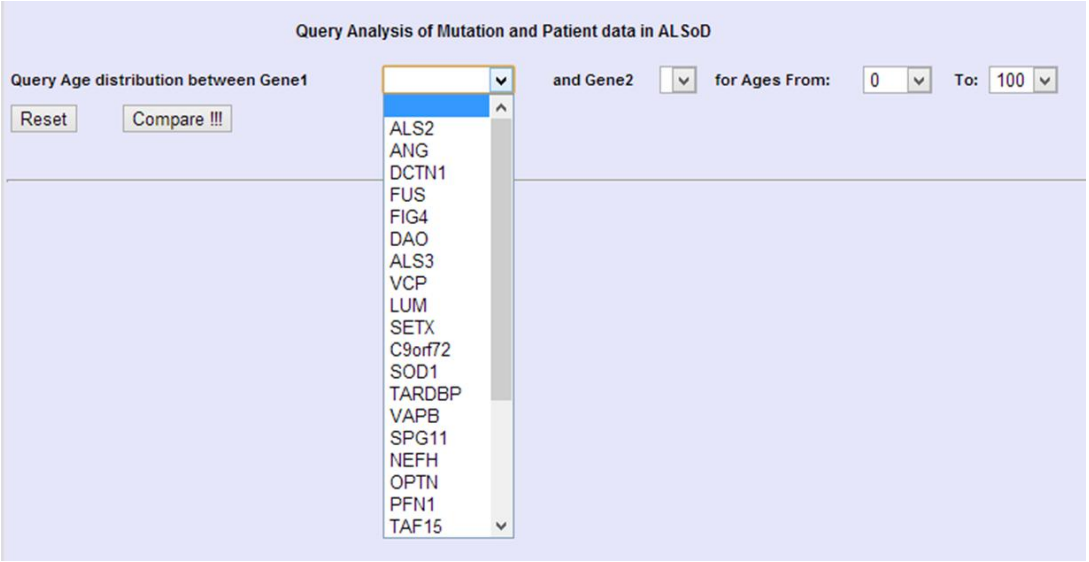


Figure 55: Side by side comparison submenu

2.5.) Clicking on the Predict pathogenicity submenu, a list of genes are automatically pooled from the database based on the availability of genes with at least a substitution mutation.

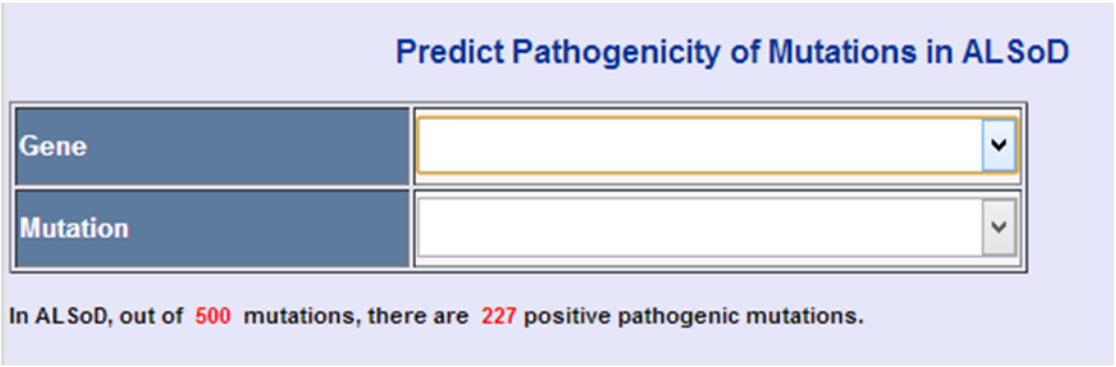


Figure 56: Pathogenicity prediction submenu

Once the gene is selected, a filtered list of mutations for the gene is displayed in the second dropdown list. Two tables could appear on the same page if there are patient data available for that mutation else one table appears. For example, a user wishes to know if an A4V mutation in SOD1 is pathogenic or not. Select the SOD1 gene and the A4V mutation. A single row table displays and combines predictions from 3 freely available online bioinformatics tools: PANTHER, PolyPhen and SIFT. The Pathogenic result is determined by combining the predictions of each tool and rating them in binary form. The result column allows a 'Yes'

prediction if at least one of the bioinformatics tools gives a positive prediction '1' else it is 'No'. In this case, a 'Yes' means that A4V is pathogenic. A tabular list of patients affected with the A4V mutation from replicated studies with references to publications are displayed also. Clicking on the VIEW CHARTS button shows. Pie charts of Gender ratio, Family ratio and Site of Onset ratio; a column chart of the mean age of onset for the selected mutation; a box plot of the mean age of onset, a doughnut chart of the country ratio where the affected patients originates from or lives are displayed.

2.6.) Clicking on the Gene credibility submenu, 11 checkboxes appear as user-configurable criteria.

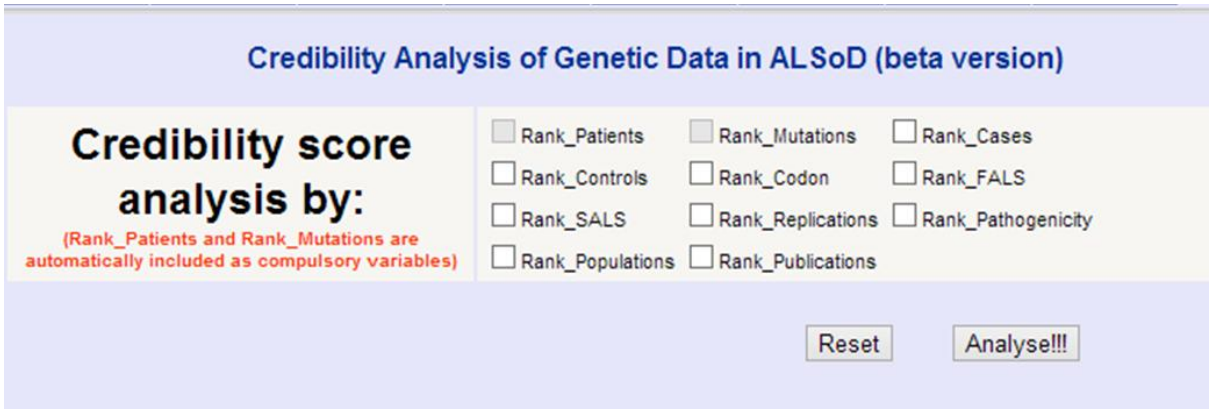


Figure 57: Credibility score analysis submenu

For example, a user needs to know the top 5 genes based on the number of pathogenic mutations and the number of countries mutations are found. The number of patients recorded and the number of mutations in a gene are selected by default and so it is greyed out. The user selects Rank_Pathogenicity and Rank_Populations checkboxes and clicks on the ANALYSE button. All the 14 FALS genes are ranked accordingly. The second section showing the details of the ranked credibility data are displayed on the same page. The user is able to click on 'NUMBER OF PREDICTED PATHOGENIC MUTATIONS BY RANK' to view the full list of pathogenic mutations and the 'NUMBER OF UNIQUE COUNTRIES ON GENES' to view the full list and total number of unique countries where patients with mutations are from. A more elaborate explanation on credibility analysis is given later in section 6.4.

6.3.3.5 Summary webpage – A cumulative report and general overview of links to sections of the database

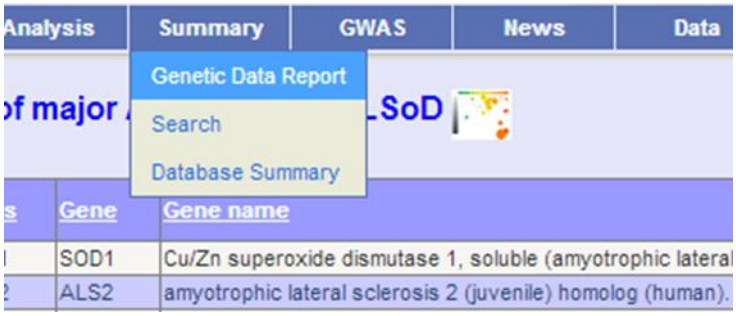


Figure 58: Summary menu

3.1) Genetic data report submenu displays the summary report of mutation and patient data on the database in 4 sections: One is a line graph of Average age of onset versus genes, a table summarizing the genetic data on each gene, a single row table showing the total number of patients with gender ratio, site of onset ratio and family ratio and finally, the top 20 most frequent mutations recorded.

3.2) Search submenu displays 2 sections for a user to easily search the database for genes and mutations either by individual gene or gene effects. For example, a user wants to find out the FALS genes found in SALS genes and details of a certain mutation V2053M in SPG11 gene.

Click on gene effect dropdownlist under the second section, select FALS genes found in SALS and this filters the gene list on the second dropdownlist. A list of all FALS genes found in SALS genes are shown. To further find the mutation details required, select SPG11 and this filters the list of all SPG11 mutations recorded on the database. Select the Val2053Met mutation, ignore the mutation type as this is not necessary for this mutation and click on the SUBMIT button.

The user is taken to a different page showing the summary of the mutation and a SELECT hyperlink which when clicked, displays mutation details, patients with this specific mutation and administrative/institution details of the mutation.

3.3) Database summary submenu displays the graphical frequency of mutations by genes and the frequency of patients by genes in bar charts. A second section also displays quick links to webpages like case summary, mutation data, patient data, data tables, last 10 mutations and patients submitted, statistics of visitors, disease information, literature, contributors, Frequently asked questions, useful external resource links and how to contact the administrators of the database.

For example, a researcher wishes to find out the last mutation and patient data added to the database. Click on cell 5 link, LAST 10 MUTATIONS & PATIENTS. Two tables are displayed and the user sees details of the last 10 mutations and the last 10 patients added to the database.

6.3.3.6 GWAS webpage - Meta-analysis and on-the-fly analysis of Genome Wide Association Studies

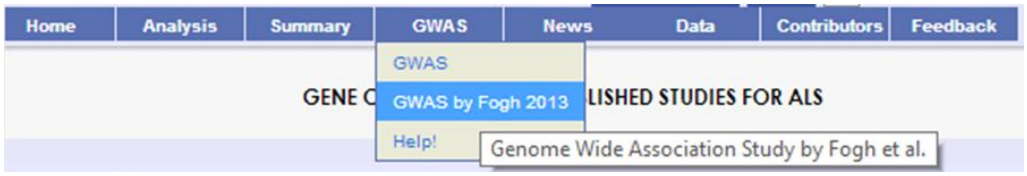


Figure 59: GWAS menu

4.1) This webpage allows users to observe the graphical overview of all submitted GWAS results in ALS, to perform an on-the-fly meta-analysis of uploaded data from users, to search for a particular SNP which could be of interest to the researcher, to view a chromosome in UCSC genome browser, to view all GWAS publications, to view all complementary GWAS websites associated with ALS and a more detailed list of published GWAS data in ALS extracted from GWAS Catalogue [543].

4.2) For example, a researcher has a list of SNPs to analyse and find out how significant these SNPs are in relation to data available on the database.

4.3) Apart from a username and password which is given to every user after registration, a text file containing the user's data is also required in the correct format.

4.4) The correct and acceptable format for the text file is the plink format but without headings. There MUST be four (4) columns: CHROMOSOME, SNP, POSITION, P-VALUE. Each field or column should be seperated with TABs. After the P-Value, there should only be return to the next line and not another TAB. On the last row, the cursor must stop at the end of the P-value column. Column specification: CHROMOSOME – Integer, SNP - String (Alphanumeric), POSITION – Integer, P-VALUE - Float (with decimals). You can export file from plink but remember to edit the text file by removing the first line (heading). The text file must be with a .txt extension and nothing else.

4.5) There are four simple steps to analyse your GWAS data in ALSoD Figures 60-62

1) Upload data - to insert text file to database,

2) Run Analysis - to run scripts,

3) View Graph - to display chart of the combined p-values

4) Chromosome on UCSC Genome Browser with track codes in Appendix 29 and 30.

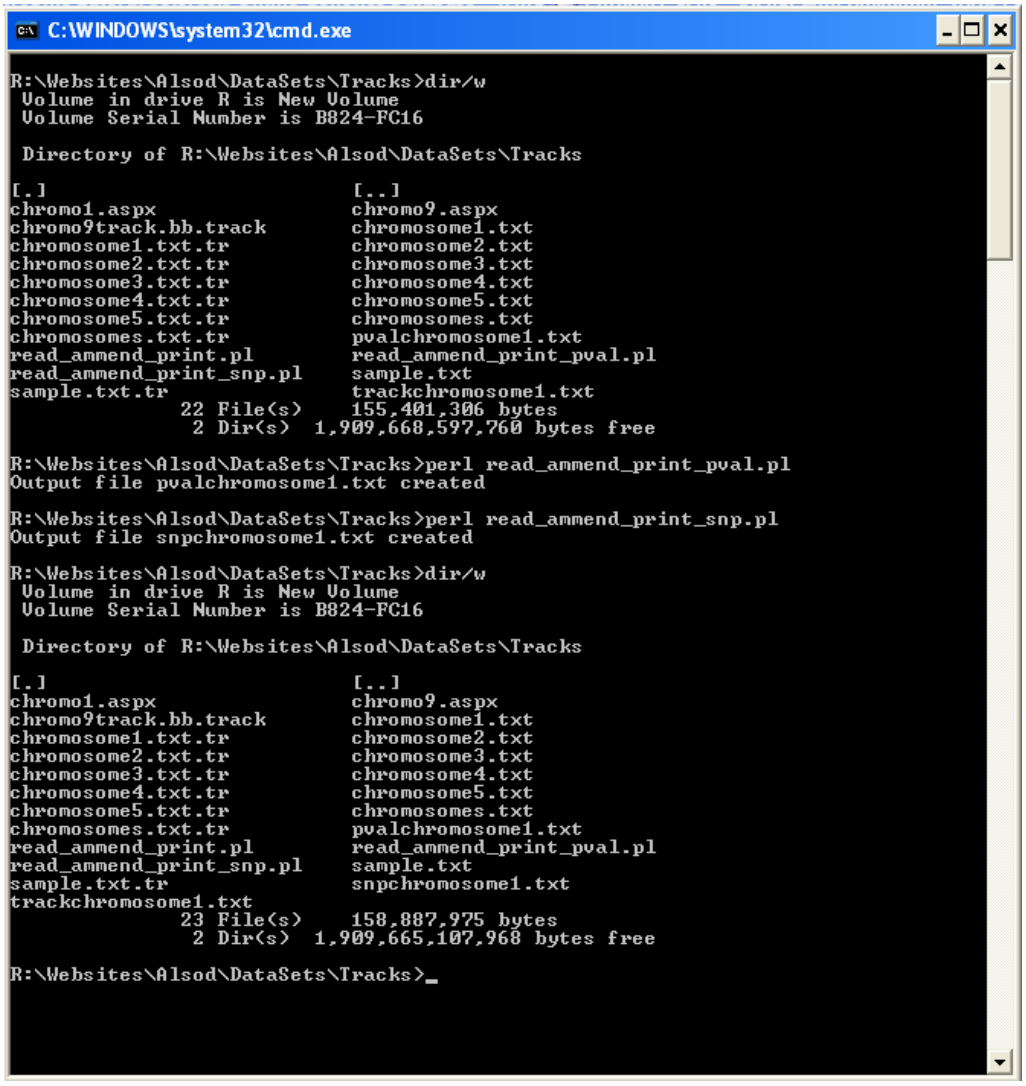


Figure 60: Customizing UCSC tracks process 1

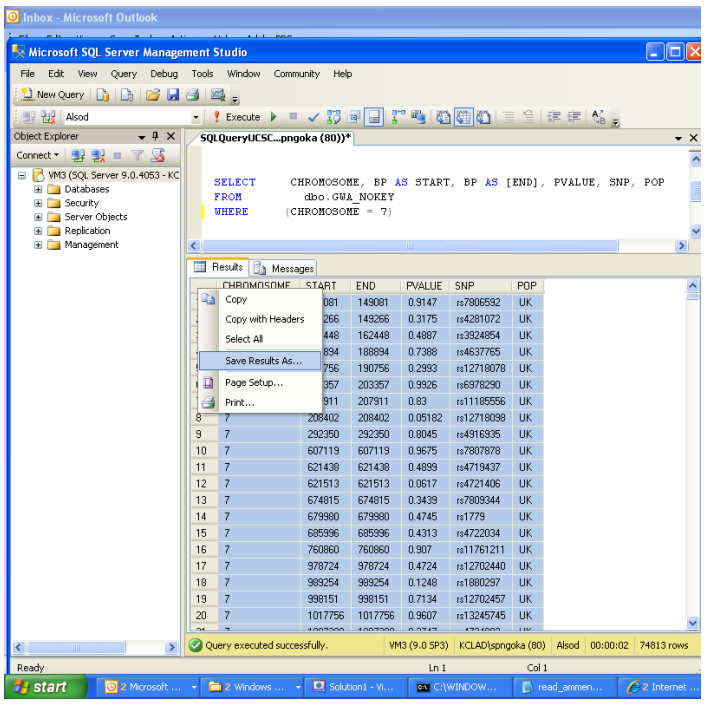


Figure 61: Customizing UCSC tracks process 2

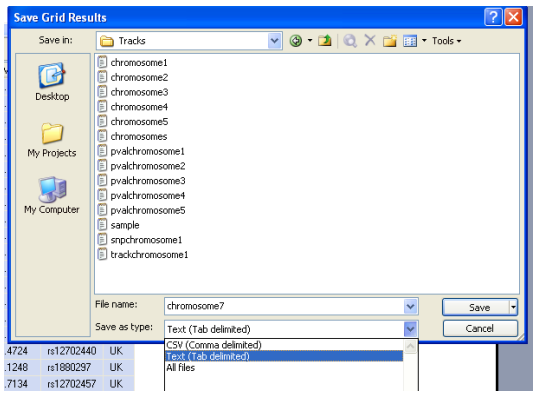


Figure 62: Customizing UCSC tracks process 3

4.6) GWAS by Fogh 2013 submenu is a meta-analysis of 8 independent studies from the latest and largest genome wide association study in ALS [544].

4.7) The Help submenu helps the user on steps to take to analyze their data.

6.3.3.7 News webpage - Latest News, blogs, books and web info on ALS

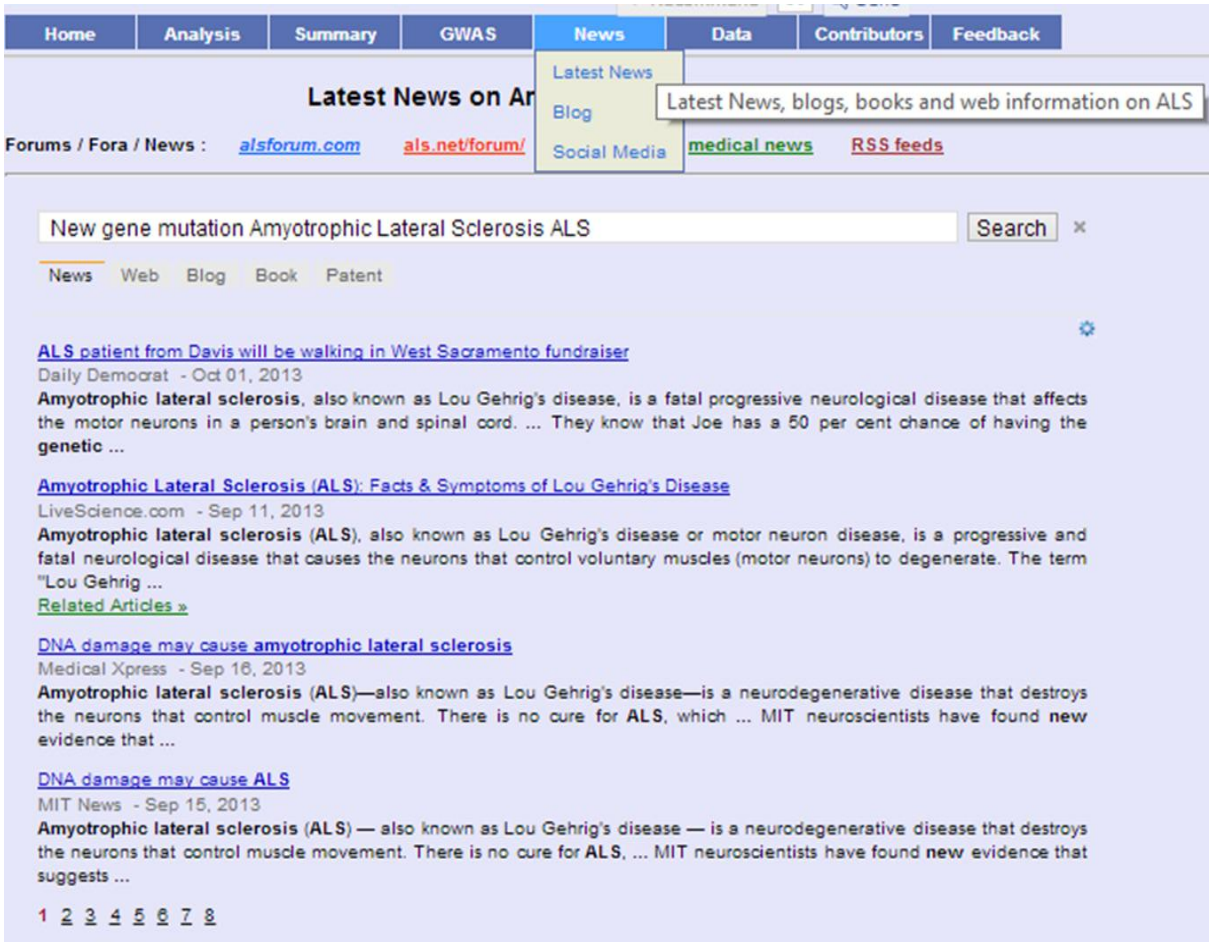


Figure 63: News menu

5.1) There are 3 submenus for the News menu: Latest News, Blog and Social Media.

5.2) The latest news displays current information in the press using RSS feeds. It is up-to-date with information on genes, mutations in ALS.

5.3) I developed a blog for users to add comments on the BlogList.aspx web page and for comments to be viewed on the BlogEntry.aspx web page. The title of a discussion, the content of the discussion, the name of user posting the comment which are filled in a form are viewed. The date and time of submitting a comment are automatically included.

The data set and structure are stored and retrieved from an XML file having four columns: Name, Time, Title and Blog [545]. Once the blog form is submitted, a new blog is updated into the XML file immediately and appears on the blog page.

5.4) A social media page in 2 columns with one column for facebook and the other for twitter.

6.3.3.8 Data webpage – All mutation and patient data available for download

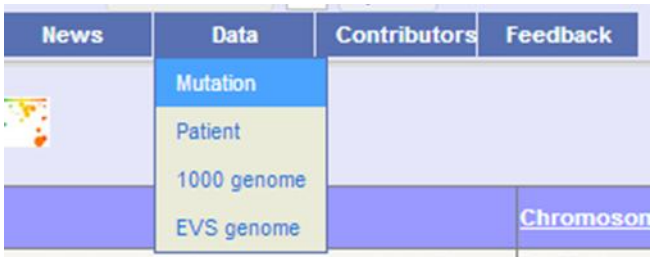


Figure 64: Data menu

6.1) The data access page shows in tabular formats full lists of all mutation and patient data submitted to ALSoD. A very useful feature is the download button 'HERE'. For example, a user wants to have the list of all mutation data on the database to find out if the mutations found in their laboratory has been previously discovered with reference.

6.2) Click on the download button on the mutation list section, a file download window pops up asking to either open or save or cancel the spreadsheet file.

6.3) An excel list of the Mutation name, Mutation code, Gene, Type, Original Amino acid, Mutated Amino Acid, Codon and mutant structure as seen on the webpage.

6.4) The mutant structure column has a SELECT hyperlink to an external database, www.bioinf.org.uk (Dr. Andrew C.R. Martin's Group) at UCL. This link is only partly developed for the SOD1 gene which was the initial focus of the database.

6.3.3.9 Contributors webpage - Names of institutions around the world collaborating with the ALSoD team on data submission

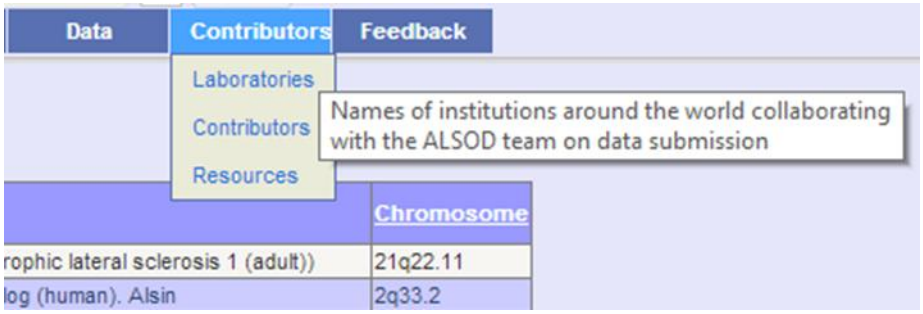


Figure 65: Contributors menu

7.1) Laboratories submenu displays a table of 7-columns (Laboratory, Address, Country, Expert, Email, Profile, Lab details). This shows a list of some major laboratories where ALS researchers conduct experiments and generate massive data.

7.2) Contributors submenu lists the names of institutions where researchers who have submitted data to ALSoD are based.

7.3) Resources submenu shows the hyperlinked logos of websites that have collaborations with ALSoD. ALSGene [539], ALS mutation database , HGMD[546], Uniprot[547], FALS Connect, PANTHER[548], SIFT[549], PolyPhen[550], PubMed, Google scholar and Jalview[551].

6.3.3.10 Feedback webpage – Interacting with users of the database

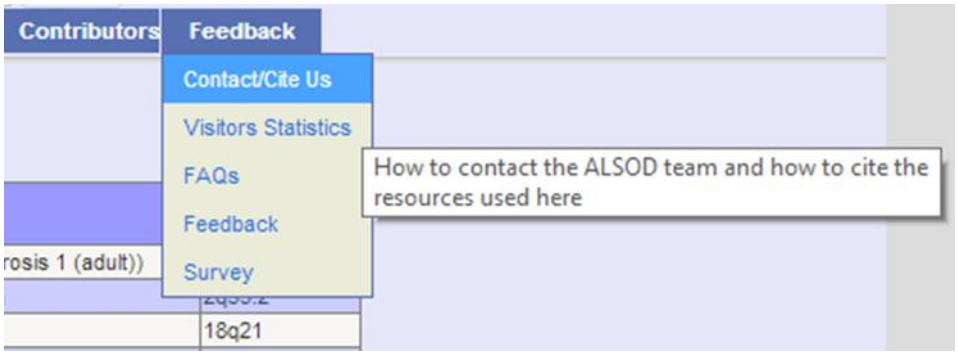


Figure 66: Feedback menu

8.1) Contact and Cite us submenu gives a brief description of where and whom to send queries, comments or suggestions to. It also describes the various versions and links to publications of ALSoD.

8.2) Visitors Statistics submenu shows a map of the world and the frequency of visits by visitors from different parts of the world. The second graph shows a column chart of country by frequency statistics since 3rd January 2009 (when the database table for the page was developed) and the total number of countries. The third graph is also a column chart of the webpages visited by the frequency of visitors. The fourth graph shows the number of unique users who have visited the website while the fifth graph shows the column chart of country by frequency of visitors who visited the website a day before. A daily visit scatter chart is the sixth graph while the last section on the page displays hyperlinks to where the ALSoD website has been mentioned online through Google search script embedded on the webpage. This helps the administrators to establish the usefulness of the website.

8.3) FAQs submenu displays top 15 frequently asked questions by users of the database.

8.4) Feedback submenu is a guestbook for visitors to fill-in some information and their candid comments. To avoid injection of automatic spams into the database by hackers, a recaptcha tool which brings up two random words required before a comment can be submitted.

8.5) Survey submenu displays the latest survey conducted by the ALSoD team.

8.) Feedback webpage – Interacting with users of the database

6.3.4 On-the-fly Meta-Analysis of Genome Wide Association Studies in ALS

6.3.4.1 Description of data

From a recent study in 2010 by Shatunov [403], we got data of a total of 288,357 SNPs used for genome wide analysis screened in a set of 1,821 sporadic ALS cases and 2,258 controls from the U.S. and Europe. Datasets each containing Chromosome, SNP, Base Pair, Amino Acid1, Frequency_in_Affected, Frequency_in_Unaffected, Amino Acid2, Chi-square, P-Value and Odd Ratio were acquired from five different sources. These sources are National Institute of Health (USA), France, Holland, United Kingdom and Boston were combined for meta-analysis as seen in Table 5.

6.3.4.2 Flowchart of GWAS on ALSoD

At the time of developing this section of ALSoD, a paper was published on protecting the confidentiality of patient data in publications [524]. I therefore had to put some features in place to safeguard the identity of patients whose data are included on the database. One of the features was to allow only researchers who login with their username and password to view the data. So, the flowchart of using GWAS on ALSoD is seen below in Figure 67.

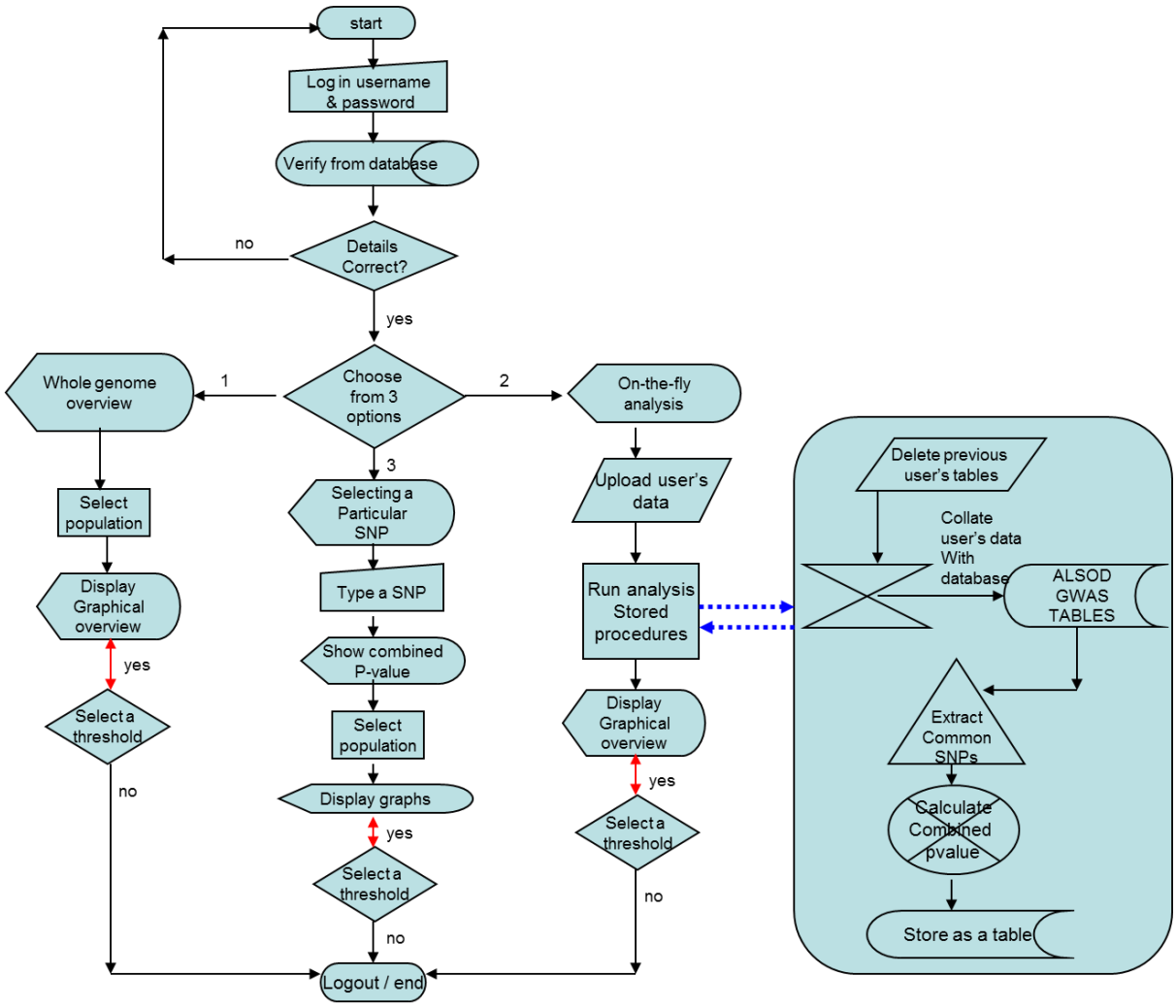


Figure 67: GWAS flowchart

6.3.4.3 Meta-Analysis

The ALSoD database schema was restructured, table designs were changed, queries were rewritten and appropriate stored procedures were implemented to allow room for expansion.

The available methods of testing for differential gene expression are not as hopeful as combining p values. Also, the combined p method is above all appropriate as a problem-solving tool for exploratory analysis of microarray data and for combining scores of a biological sequence (DNA or protein). P-Values of datasets were combined using the Fisher's method by clicking on the 'Fisher Combine P-Values' button in Haploview. That is, using the traditional chi-square test, If there are 'x' independent measurements, the resulting p-values will be independent and then the sum of x such p-values as (-2log p1) + (-2log p2) + ... + (-2log px) will follow chi-square distribution with 2x degree of freedom.7

I plotted an X-Y axis graph where the X axis is for chromosomes and the Y axis is –log10 scale of the combined P-Values. This is saved and used to verify the dynamically generated graph as seen in Figure 1. I later exported the PLINK format result data to text files. These text files were inserted into the SQL database using a bulk-insert stored procedure.

6.3.4.4 GWAS tables

The steps below will show how the tables and views used for the genome wide association data analysis were derived.

Five text files of rows and columns separated by tabs were copied on a USB stick from my supervisor. These files are for UK, Boston, Holland, NIH and France.

These individual files were stored in a newly created folder on the web server called 'DataSets'. The first row containing the headers was removed and the filename was changed to avoid conflict. An example of a boston text file named assoc_bos.txt (with header) was changed to bos_result_assoc.txt (without header) and stored in `\\vm3\Alsod\DataSets\bos\bos_result_assoc.txt` as shown below:

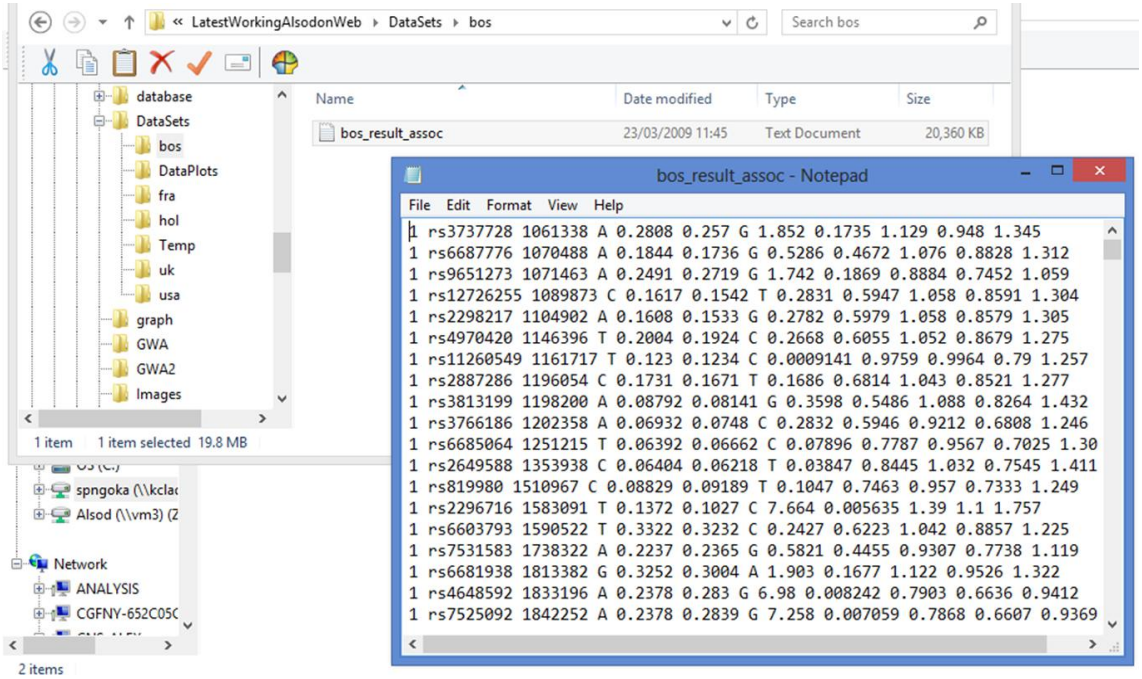


Figure 68: GWAS text file stored on SQL database

On SQL database, a table for boston data was created and called GWA_BOS as shown:

```
CREATE TABLE dbo.[GWA_BOS](
        [CHROMOSOME] [int] NULL,
        [SNP] [varchar](30) PRIMARY KEY,
```

```
    [BP] [int] NULL,
    [A1] [varchar](1) NULL,
    [F_A] [float] NULL,
    [F_U] [float] NULL,
    [A2] [varchar](1) NULL,
    [CHISQ] [float] NULL,
    [PVALUE] [float] NULL,
    [ODDRATIO] [float] NULL,
    [L95] [float] NULL,
    [U95] [float] NULL)
GO
```

The text file was uploaded into the GWA_BOS table using the 'BULK INSERT' function:

```
BULK INSERT dbo.[GWA_BOS]
FROM '\\vm3\Alsod\DataSets\bos\bos_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO
```

The steps above are repeated for GWA_UK, GWA_FRA, GWA_HOL and GWA_USA using uk_result_assoc.txt, fra_result_assoc.txt, hol_result_assoc.txt and usa_result_assoc.txt as text files respectively.

The GWA_NOKEY is the single table created to house all five individual population text files by uploading the files one at a time into the same table as shown in Appendix 15.

I took note of the number of rows uploaded each time a population text file was uploaded to GWA_NOKEY table and this table was modified using the 'ALTER TABLE' function as shown in Appendix 16.

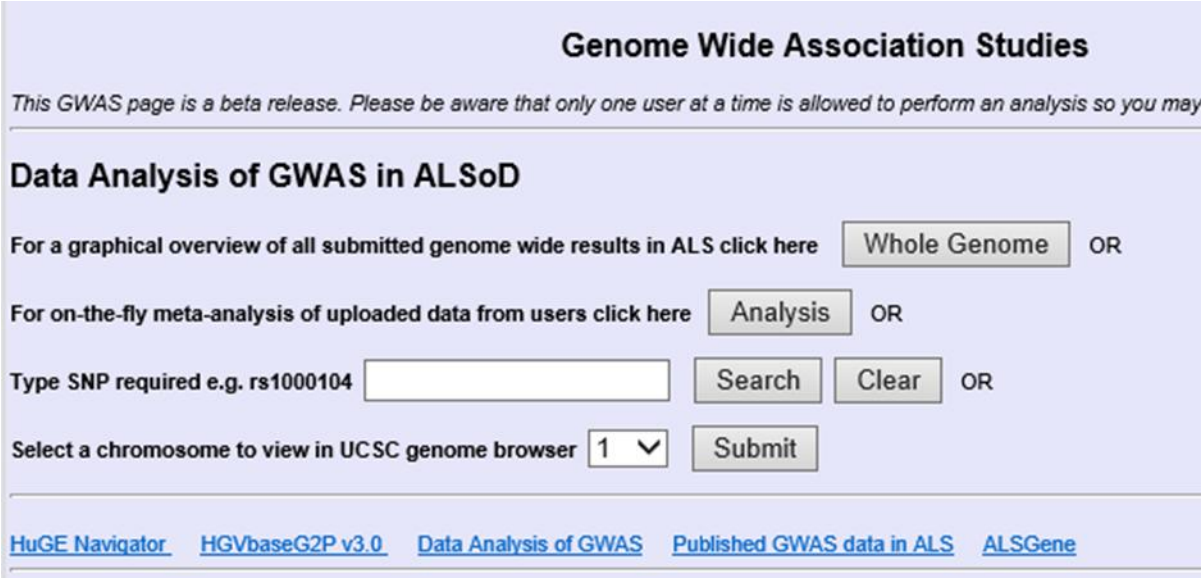


Figure 69: Genome Wide Association Study webpage

A user is taken to the page above where 4 tasks could be carried out as seen in Figure 69.

6.3.4.5 Whole Genome

For a graphical overview of all submitted genome wide results in ALS click on 'Whole Genome' button

A list of checkboxes for 5 populations (UK, Boston, Holland, France, NIH) are displayed on the screen. A user is asked to select population(s) to view graph and click the submit button as shown below in Figure 70.

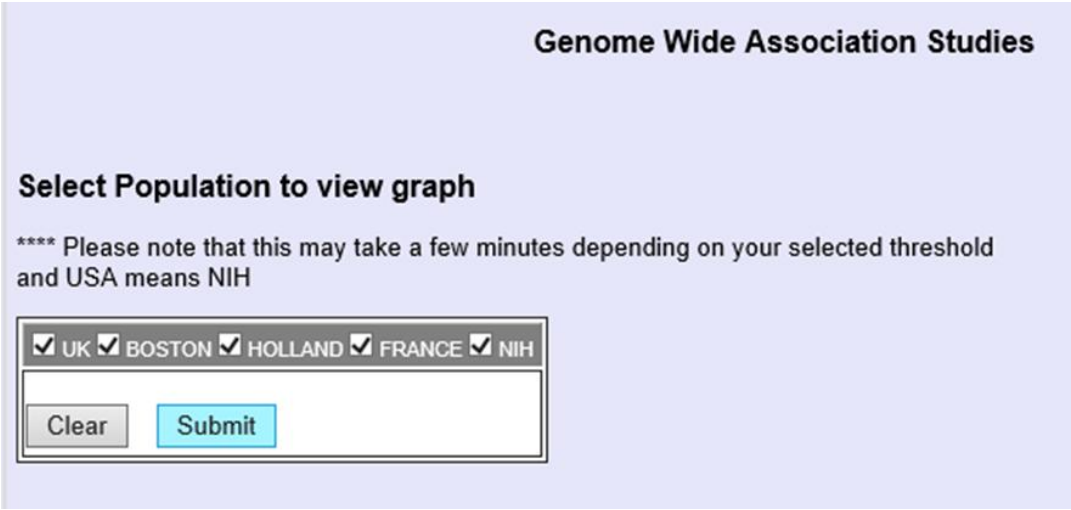


Figure 70: Whole Genome analysis

6.3.4.5.1.1 Populations and combinations

The database table 'GWA_NOKEY' –already described above - contains 1,420,350 rows on the VM3 server and it takes 46 seconds to retrieve the table as seen below in the yellow status bar:

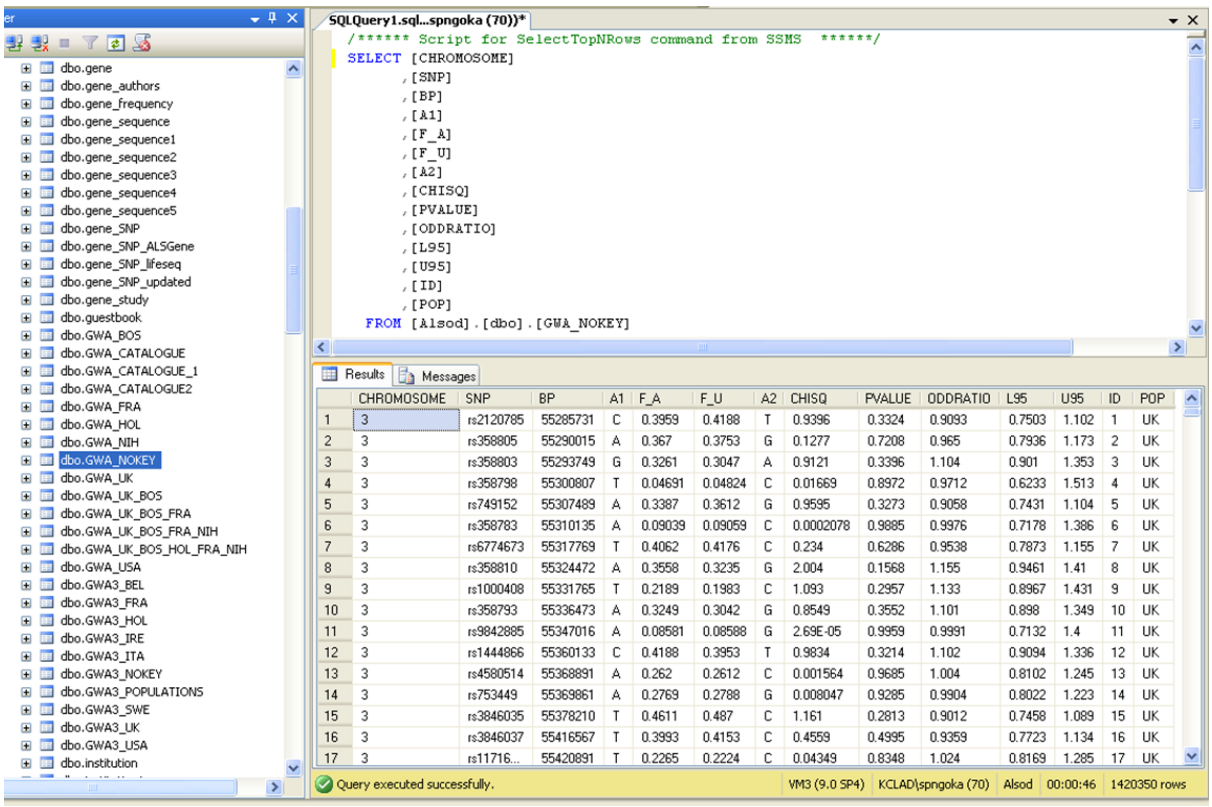


Figure 71: GWA_NOKEY table

This GWA_NOKEY table contains all the genome wide data for the 5 populations and POP column distinguishes the name of the population where:

UK is United kingdom

BOSTON is Boston

HOLLAND is Netherlands

FRANCE is France

NIH is National Institutes of Health in USA

GWA_NOKEY is queried using:

SqlDataSource2 = "SELECT DISTINCT [POP] FROM dbo.[GWA_NOKEY]"

A list of checkboxes are displayed according to the generated number of populations in the database using

<asp:CheckBoxList ID="CheckBoxList1" runat="server" DataSourceID="SqlDataSource2" DataTextField="POP" DataValueField="POP" RepeatColumns="5" RepeatDirection="Horizontal" CellSpacing="5" OnSelectedIndexChanged="viewChart_Click" />

To concatenate the selected values of the checkboxes I used:

```
For Each item As ListItem In CheckBoxList1.Items
    If (item.Selected = "true") Then
    Label3.Text += item.Text
    End If
Next
```

With **n** populations, the possible combinations in **r** ways are derived using the formula:

Equation 1 : Combination formula

$$\frac{n!}{(n-r)!\,r!}$$

To choose 5 combinations out of 5 countries is 5!/ (5-5)!5! = 1/1 = 1

To choose 4 combinations out of 5 countries is 5!/ (5-4)!4! = 5/1 = 5

To choose 3 combinations out of 5 countries is 5!/ (5-3)!3! = 20/2 = 10

To choose 2 combinations out of 5 countries is 5!/ (5-2)!2! = 20/2 = 10

To choose 1 combination out of 5 countries is 5!/ (5-1)!1! = 5/1 = 5

Total = 1 + 5 + 10 + 10 + 5 = **31 possible combinations**

Meanwhile, I downloaded the haploview.jar software on the primary drive of my machine. This software enabled me to combine the 5 populations into 31 combinations of tables and to plot graphs.

6.3.4.5.1.2 Haploview-generated result

A program called ‘Haploview’ developed in the lab of Mark Daly at the Broad Institute of MIT and Harvard was designed to analyze association study tasks on HapMap data, choosing tag-SNPs, evaluate the value of genotype data for a disease, investigate for association, and evaluate the area of a chromosome with a positive association. [552, 553]

Using Haploview software downloaded from http://www.broad.mit.edu/mpg/haploview/ on my PC, an image was generated and coded on the webpage http://alsod.iop.kcl.ac.uk/GWA2/whole.aspx. Steps below were used to generate the haploview image.

6.3.4.5.1.2.1 Installing Haploview software

From the Broad Institute website, I followed the instruction by installing Java software on my computer first by clicking on ‘download Java’ and the process continued.

Figure 72: Downloading java software

I Installed haploview.jar file on a preferred location (desktop) on my machine. Haploview opens directly using java if java is installed first else, it would ask the user how the file should be opened if there is no java on their machine.

6.3.4.5.1.2.2 Uploading data on Haploview

Double-click on haploview.jar file and a welcome screen comes up as shown in Figure 73.

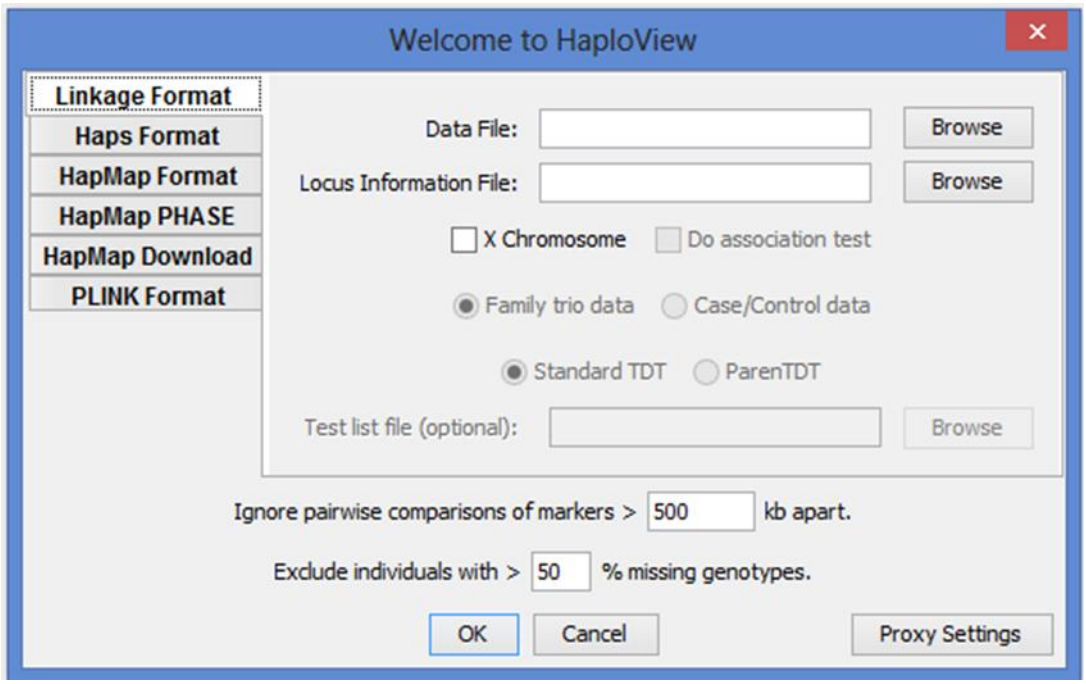


Figure 73: Haploview interface

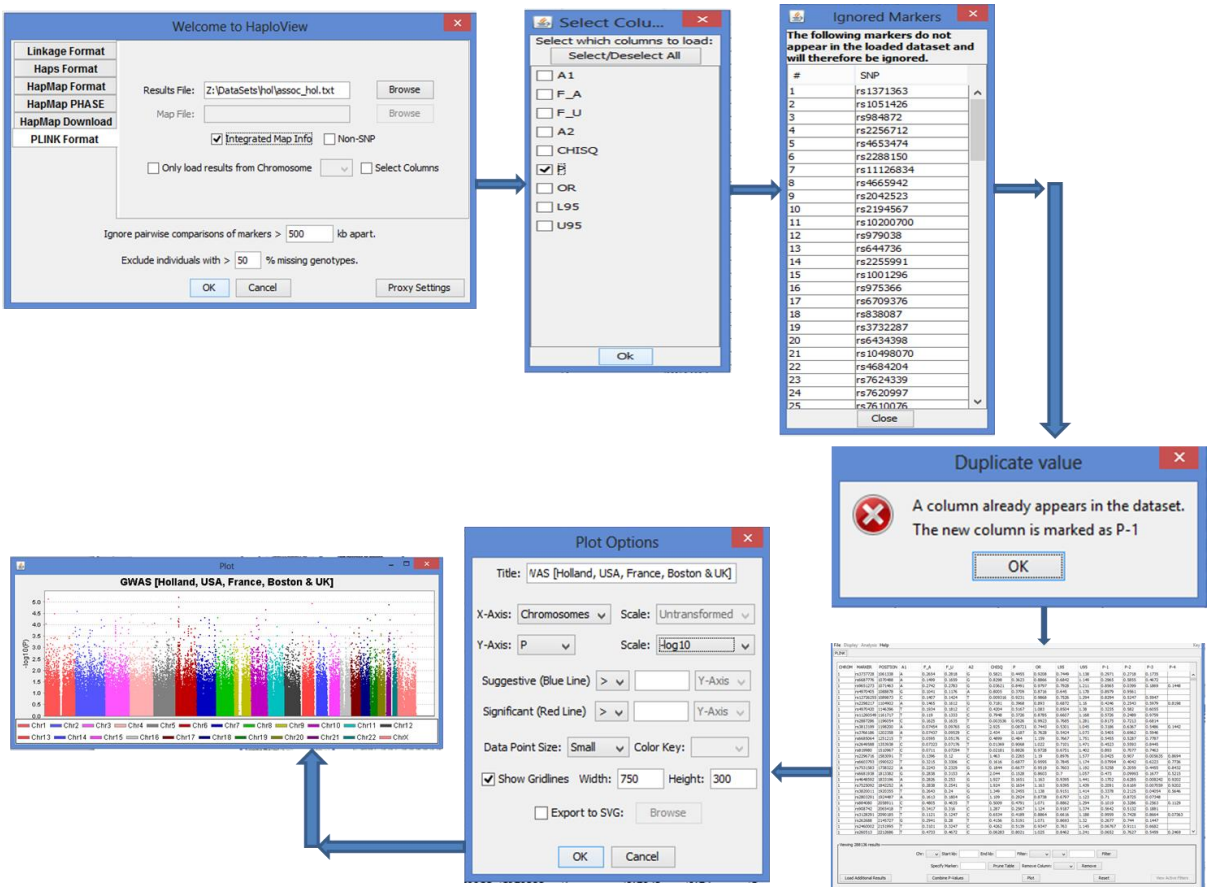
The PLINK format embedded in the Haploview software (http://www.broadinstitute.org/mpg/haploview) was used to combine the data into groups by identifying the common SNPs from various combinations of datasets in an organized manner starting with the Holland population as seen in Table 5: GWAS of 5 populations below.

If combining two populations, open the bigger text file first and then the smaller one e.g to combine UK and Boston, open the Boston file first as it has 20,360KB size while UK is uploaded next.

Table 5: GWAS of 5 populations

S/N	Abbreviation	Population	File size (KB)	No of rows
1	UK	United Kingdom	19,873	275,619
2	BOSTON	Boston	20,360	282,448
3	HOLLAND	Netherlands	20,767	288,136
4	FRANCE	France	20,650	286,507
5	NIH	United States	20,734	287,640

Use the ‘Load Additional Results’ button to add more data until the 5 populations are included. The order of data upload is Holland, NIH, France, Boston, UK.



6.3.4.5.1.2.3 Attaching image to webpage

The possible graphs of 31 combinations were plotted and saved on the database with their corresponding names as seen in Appendix 25.

The image is attached to the webpage using:

<asp:Image ID="Image4" runat="server" Width="750px" Height="300px" />

6.3.4.5.1.2.4 Code-generated result

The generated chart is embedded on the home page as the dynamically generated version of the haploview static image.

A group of radio buttons allows users to choose the p-value threshold from >=2 to >=5

```
<asp:RadioButtonList ID="RadioButtonList1" runat="server" Height="95px" Width="67px" RepeatColumns="1"
CellSpacing="5" RepeatDirection="Vertical"                OnSelectedIndexChanged="viewChart_Click">
  <asp:ListItem Value="0"> >=2</asp:ListItem>
  <asp:ListItem Selected="True" Value="1"> >=3</asp:ListItem>
  <asp:ListItem Value="2"> >=4</asp:ListItem>
  <asp:ListItem Value="3"> >= 5</asp:ListItem>
  </asp:RadioButtonList>
```

A label is assigned to the selected radio button

```
Label4.Text = RadioButtonList1.SelectedValue
```

The Chart is configured to display –log10 p-values for the chosen population combinations. By default, the negative log10 p-values are set to start from 3.0 or greater

```
        If Label4.Text = "0" Then
      query = "log_ukbos >= 2"
      Chart1.Titles.Add("GWA P-VALUES [UK, BOSTON] THRESHOLD >= 2")
    ElseIf Label4.Text = "1" Then
      query = "log_ukbos >= 3"
      Chart1.Titles.Add("GWA P-VALUES [UK, BOSTON] THRESHOLD >= 3")
    ElseIf Label4.Text = "2" Then
      query = "log_ukbos >= 4"
      Chart1.Titles.Add("GWA P-VALUES [UK, BOSTON] THRESHOLD >= 4")
    ElseIf Label4.Text = "3" Then
      query = "log_ukbos >= 5"
      Chart1.Titles.Add("GWA P-VALUES [UK, BOSTON] THRESHOLD >= 5")
```

End If

X- and Y-axis values of the chart are derived from the database depending on the population(s) selected. Using the UK and Boston example, the database query reads:

"SELECT TOP (100) PERCENT chr, log_ukbos, snp FROM dbo.[GWA_UK_BOS] where " + query + " ORDER BY chr"

I attached the X and Y values to the chart area for each of the 31 combinations:

Chart1.Series("Gwa").XValueMember = "snp"

Chart1.Series("Gwa").YValueMembers = "neglogpval"

6.3.4.5.2 On-the-fly Analysis

Stored procedures were written to clear old tables from the database used for analysis by a previous user, to verify the format of the user's data and to upload it to the database as a new table.

I then wrote some T-Sql queries to compare the existing tables with the newly created table made up of the user's data. New p values are generated by using the fisher's method as described above. That is,

Equation 2: Fisher's method

$$P\ Value = Chi\text{-}sq\ [\{(-2\log p1) + (-2\log p2) + \ldots\ldots\ldots + (-2\log p6)\}\ ,\ 12\ d.f.]$$

The Y axis (negative log of the new P value) is plotted against the chromosomes on the X axis.

A chi-squared distribution dialogue box is also included in the tool for verifying the chi-square and natural log values where the cumulative p value and the degree of freedom are known as seen above.

6.3.4.5.3 Search for SNP id

The search button links to a webpage http://alsod.iop.kcl.ac.uk/GWA2/gwa.aspx. A SNP id is typed into the textbox e.g. rs1025542 and displays a table with the population, p-value and the combined p-value; a non-clickable static graph generated from haploview, a clickable dynamic graph generated in realtime, the graphical position of the SNP on the chromosome and it's base pair location. Codes are in Appendix 31.

6.3.4.5.4 Chromosomal view on UCSC genome browser

On UCSC genome browser, http://genome.ucsc.edu/, I customized the dataset of 5 populations and embedded the web address into ALSoD website using <Iframe>. For example, using chromosome 1 dataset, follow the steps below:

Step 1 – Open the UCSC web address, click on 'Genome Browser' menu, click on 'add custom tracks', add the track file in form of web address 'http://alsod.iop.kcl.ac.uk/DataSets/Tracks/resultchromo1track.txt' to the textbox, click on 'go to genome browser' and after a few seconds, the result is displayed as shown in Figure 74, Appendices 29 and 30.

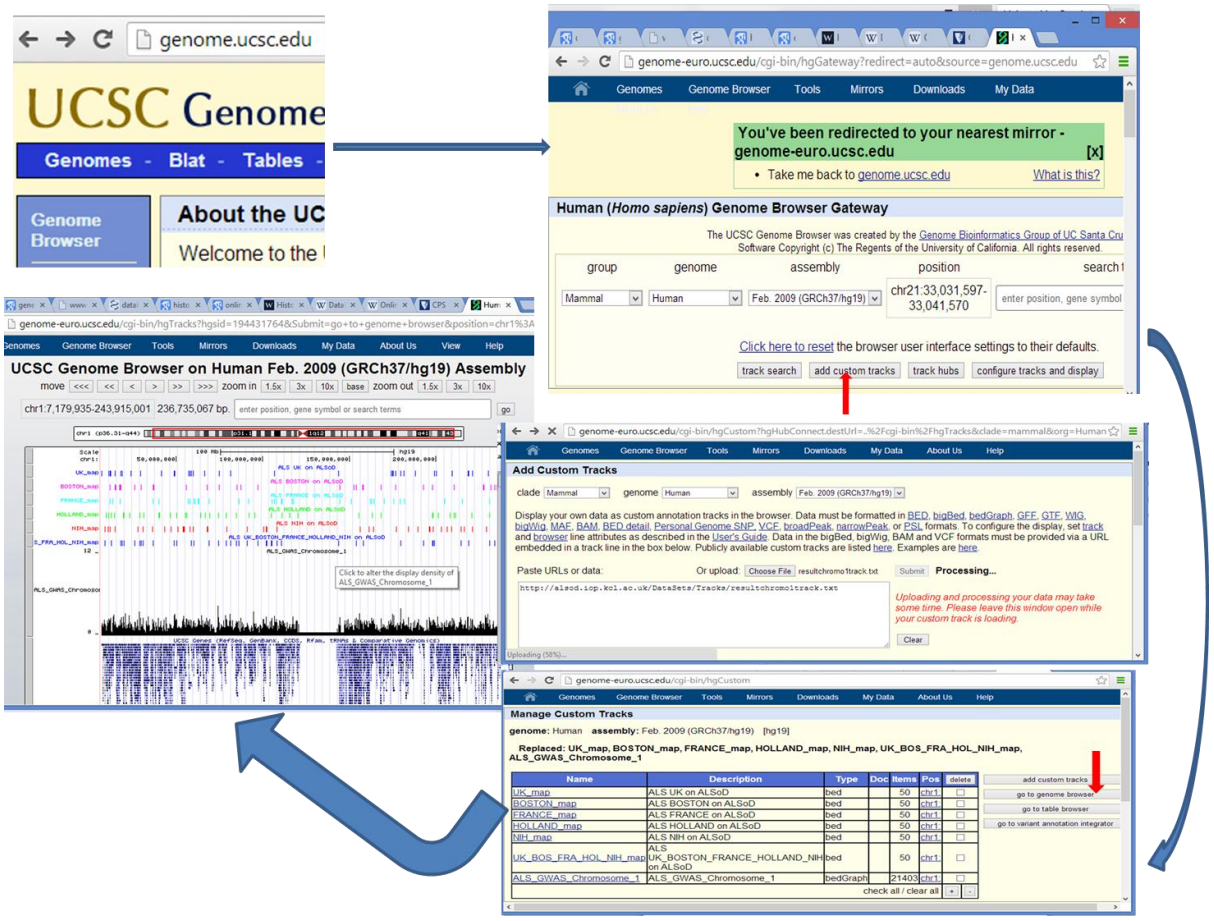


Figure 74 : UCSC genome browser track configuration

Step 2 – Dynamically analyse and embed image generated automatically on the webpage http://alsod.iop.kcl.ac.uk/DataSets/Tracks/chromo1.aspx. This image is hypelinked to the UCSC browser webpage. Below is the ASP.NET code for embedding the iframe into ALSoD.

<iframe id="Iframe1" name="I1" src= "http://genome.ucsc.edu/cgi-bin/hgTracks?org=human&position=chr1&hgt.customText=http://alsod.iop.kcl.ac.uk/DataSets/Tracks/resultchromo1track.txt" title="Chromosome 1 SNP Track as at 1st October 2010" style="width: 80%; height: 558px;">

Your browser does not support iframes.

</iframe>

6.3.4.5.5 Hyperlinks

HuGE Navigator - is an integrated, searchable knowledge base of genetic associations and human genome epidemiology.

URL: http://www.hugenavigator.net/HuGENavigator/gWAHit.do;jsessionid=E02FF44C696717C27038AFC39BB31940?query=ALS&Mysubmit=simple&geneOrderType=geneA

HGVbaseG2P v3.0 - is a comprehensive GWAS database, with powerful browser support for multi-study viewing and comparison.

URL: http://www.hgvbaseg2p.org/study/HGVST127

Published GWAS data in ALS - Data here are derived from GWAS Catalogue which is a catalogue of Genome-Wide Association Studies

URL: http://www.genome.gov/gwastudies/

ALSGene - database will provide a comprehensive, unbiased and regularly updated field synopsis of genetic association studies performed in ALS.

URL: http://www.alsgene.org/

6.3.4.5.6 Publications

Research publications containing GWAS data are listed on the website e.g.

"Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study" Lancet Neurol. 2010 Aug 27. [Epub ahead of print] Abstract

Shatunov A, Mok K, Newhouse S, Weale ME, Smith B, Vance C, Johnson L, Veldink JH, van Es MA, van den Berg LH, Robberecht W, Van Damme P, Hardiman O, Farmer AE, Lewis CM, Butler AW, Abel O, Andersen PM, Fogh I, Silani V, Chiò A, Traynor BJ, Melki J, Meininger V, Landers JE, McGuffin P, Glass JD, Pall H, Leigh PN, Hardy J, Brown RH Jr, Powell JF, Orrell RW, Morrison KE, Shaw PJ, Shaw CE, Al-Chalabi A

Each list contains the title of the publication, a pubmed link to the abstract on NCBI website and names of authors.

6.3.5 Collaborations

Because many ALS gene variants are found in both familial and apparently sporadic ALS, a two-way link out to the ALSGene database provides evidence of association to complement the genotype-phenotype correlation available from familial ALS information in ALSoD [486]. A similar link out to fALS Connect, which is a collaboration between multiple interested agencies in the US, including the patient organization The ALS Association and the research group The Northeast ALS (NEALS) Clinical Trials Consortium, makes ALSoD relevant for patients and carers as well as the scientific community. The database is adopted into the Human Variome Project (http:// www.humanvariomeproject.org) and the GWAS Phenomap Project (http://www.gwascentral.org/gwasphenomap/index.php/collaborators).

6.3.6 Embedded bioinformatics tools

ALSoD uses third party open source bioinformatics tools to embed computational analysis within the database using Java applets. For example, Figure 2: A screenshot of the Multiple Alignment and Mutations on *SOD1* gene using a combination of Claustalw and Jalview [551] is used to provide multiple sequence alignments in other species for selected genes. GeneMANIA [509] allows users to select genes of interest for prediction of interactions. A Google Earth API is used for viewing maps of mutation, risk and exposure distributions.

6.3.6.1 Claustalw and Jalview

Other species with their sequence and multiple alignments are shown for each gene although only SOD1 is the updated working version available in Appendix 18.

6.3.6.2 GeneMANIA

This interface shows other genes that interact with this gene in humans. An algorithm of how genes were selected for analysis is shown in Appendix 19.

6.3.6.3 Google Earth API

The Google Earth Plug-in with the use of JavaScript API allowed me to embed a 3D digital globe into the ALSoD website linked to corresponding genes e.g. http://alsod.iop.kcl.ac.uk/maps/mutationmap.aspx?gene_id=VAPB.

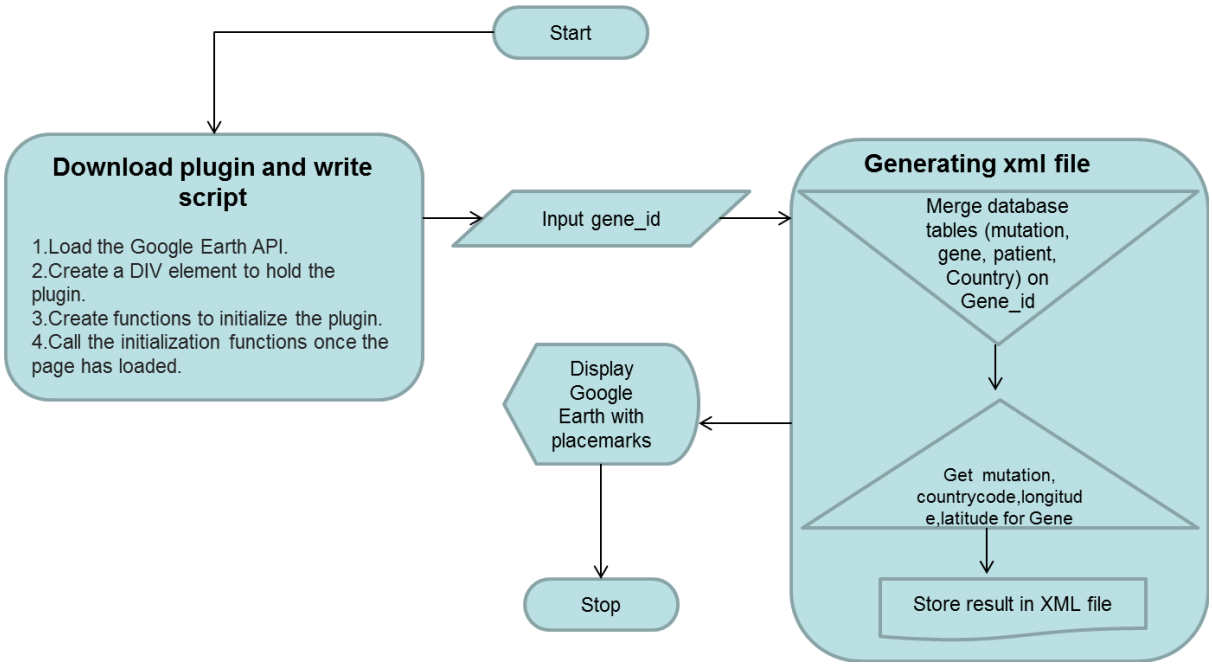


Figure 75 : Flowchart of Google Earth

The codes written on the webpage embedding the Google Earth tool are in Appendix 20.

6.3.7 Imported bioinformatics data

Data were generated from some useful external bioinformatics sites.

6.3.7.1 Haploview - www.broadinstitute.org/haploview

This was used to analyse p-values of 5 populations in the Genome Wide Analysis System producing a colourful image of SNPs graphically plotted as scattered chart with distinct colours for each chromosome.

6.3.7.2 PANTHER – http://www.pantherdb.org/tools/csnpScoreForm.jsp

In the process of finding the pathogenicity of mutations, I used the website to get the subPSEC score for every possible mutation in a gene.

6.3.7.3 SIFT - http://sift.jcvi.org/www/SIFT_BLink_submit.html

I also extracted the score from the analysis of every possible codon on a gene while finding the pathogenicity of mutations available on the database.

6.3.7.4 POLYPHEN - http://genetics.bwh.harvard.edu/pph2/index.shtml

PolyPhen-2 (Polymorphism Phenotyping v2) is a tool for predicting possible effect of an amino acid substitution on the structure and function of a human protein[554] . PSIC scores were analysed for each mutation individually to determine the pathogenicity of each mutation.

6.3.7.5 GWAS Phenomap Project – http://www.gwascentral.org/gwasphenomap

The GWAS PhenoMap project unifies the GWAS databasing research community to synchronize the phenotypic explanations used across GWAS databases with the aim of enhancing GWAS data exchange and analysis.

6.3.7.6 ALS Mutation Database -

https://reseq.lifesciencedb.jp/resequence/SearchDisease.do?targetId=1

In 2010, an ALS mutation database constructed as part of the Life Science Integrated Database Project conducted by the Japan Ministry of Education, Culture, Sports, Science, and Technology was published. It contains their original experimental results and published data extracted from scientific journals. The database is expected to play a complementary role to the ALSOD database especially in collecting variations in the Asian region [485]. I was able to retrieve some unpublished mutation and patient data although this was done with caution. This database was mainly used to populate the replicated mutation table of ALSoD as the novel mutations are already available on ALSoD.

6.3.7.7 ALSGene – http://www.alsgene.org

Meta-analysis images generated by ALSGene team from SPSS were displayed on the ALSGene website and permission was given to display these images on ALSoD under the condition that a form of reference will be made to their website. These images were displayed on gene overview page of corresponding genes were available.

'Shows ALSgene Forest plots available

If ((geneid = "APEX1") Or (geneid = "PON1") Or (geneid = "PON2") Or (geneid = "PON3") Or (geneid = "ANG") Or (geneid = "HFE") Or (geneid = "C9orf72") Or (geneid = "VEGFA") Or (geneid = "APOE") Or (geneid = "UNC13A")) Then

ImageMap1.Visible = "true"

Else

ImageMap1.Visible = "false"

End If

<asp:ImageMap ID="ImageMap1" runat="server" Visible="false"

Height="358px" Width="650px">

<asp:RectangleHotSpot AlternateText="ALSGene: A database for amyotrophic lateral sclerosis genetic association studies developed by the Max Planck Institute for Molecular Genetics Berlin, the Alzheimer Research Forum and Prize4Life. Last Update: 14th August 2011" Left="0" Top="0" Right="800" Bottom="506" NavigateUrl="http://www.alsgene.org/" Target="_blank" </asp:ImageMap>

6.3.7.8 MGI – http://www.informatics.jax.org/marker/

I downloaded a text file of 164225 rows and 10 columns from Jackson's lab into a spreadsheet. These column headers are: HomoloGene ID, Common Organism Name, EntrezGene ID, Mouse MGI ID, HGNC ID, OMIM Gene ID, Genetic Location, Genomic Coordinates (mouse: GRCm38, human: GRCh37.p10), Strand Name, Synonyms

6.3.7.9 HGVS nomenclature – http://www.ncbi.nlm.nih.gov/projects/SNP/tranSNP/tranSNP.cgi

A script written by my colleague (Simon Topp) was used to retrieve NM, NP, basepair location and rsids for mutations without these information on the ALSoD database. This was initiated as a way of improving the functionality of ALSoD and to make it more integrateable with other databases freely available online. A publication on genetic variations in reference databases [555] for analysing mutation data partly derived from ALSoD originated the idea of filling the gaps on our database so that the internationally accepted format for generating mutation names could be applied on the database. From their methods section it is clear, that after downloading the contents of ALSoD, it was necessary for them to map each variant to a genomic co-ordinate before they could perform their analysis. Hence the addition of genomic co-ordinates and dbSNP ids were generated to greatly improve the usability and functionality of ALSoD.

6.3.7.10 NCBI - http://www.ncbi.nlm.nih.gov

The National Center for Biotechnology Information develops science and health by allowing freely available access to biomedical and genomic information online. Some of the gene ids and SNP ids with other information like publication documentation are extracted from here.

6.3.7.11 HGMD - http://www.hgmd.org/

The Human Gene Mutation Database (HGMD®) represents an attempt to collate known (published) gene lesions responsible for human inherited disease. and is maintained at the Institute of Medical Genetics in Cardiff [474, 556]. This database was used at the initial stage of taking over ALSoD in 2008.

6.3.7.12 1000 genome - ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20110521/

A perl script written by a colleague (Simon Topp PhD) was used to download and extract relevant data from an external database on www.100genomes.com. This was later stored on the ALSoD database using T-Sql scripts.

6.3.7.13 EVS - http://evs.gs.washington.edu/EVS/

Here also, Exome Variant Sequence was downloaded from www.EVS.com and extracted using perl scripts. This was later stored on the ALSoD database using T-Sql scripts.

6.3.8 Population Frequency from 1000genome and EVS

A python script generated by a colleague (Kuang Lin PhD) was written and run on a text file containing mutations, their respective genes, basepair locations and HGVS nomenclature of the variants.

I downloaded python software from www.python.com, installed it on my computer and navigated to it.

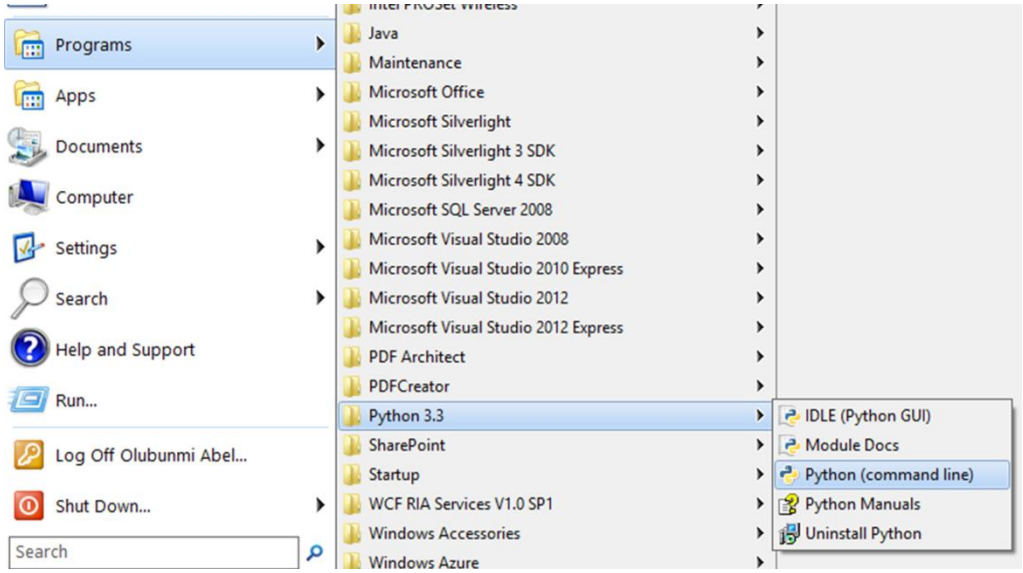


Figure 76: Navigating to Python application on desktop

From the command line, the executable file of python was navigated to and the code file executed. The code is shown in Appendix 21 while I typed “C:\Python33\python.exe snp_match.py” to execute the code as seen below in Figure 77.

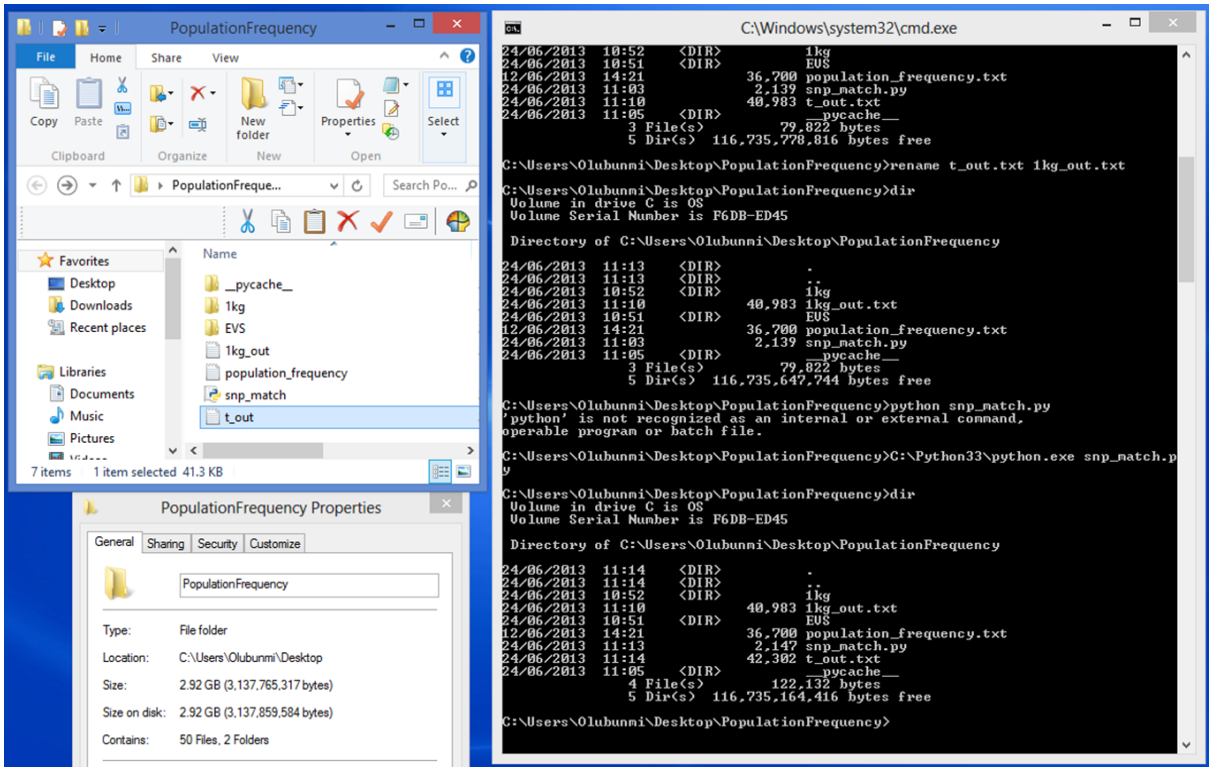


Figure 77: Screenshot of executing python

The file sent via email by my colleague was run on the command line yielding the same result as shown in Figure 78.

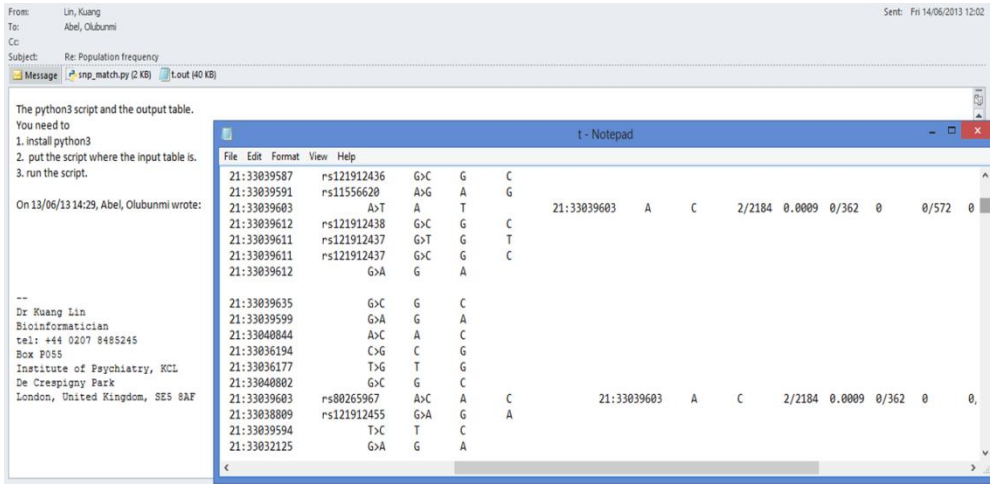


Figure 78: Result of running Python script

The output file is opened in a spreadsheet to keep the data in a proper format acceptable on the database.

Each column is separated by tabs and formatted to text.

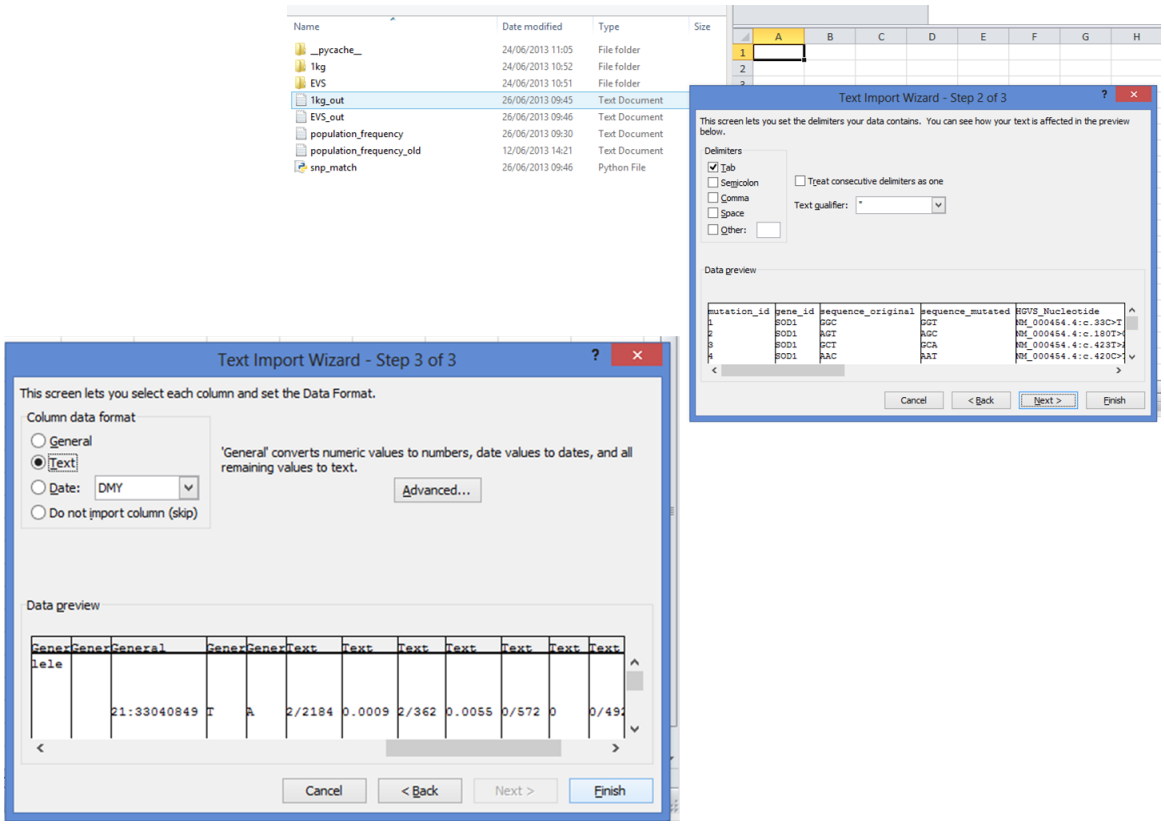


Figure 79: Screenshot of reformatting Python output in spreadsheet

A table was created to store as a text file and uploaded to the database.

6.3.9 Animal Models

Images of animals used to represent the models were extracted from Google image. These are freely available images which require no permission.

Huma (Homo Sapien) 

Cattle (Bos Taurus) 

Chimpanzee (Pan troglodytes) 

Dog (Canis lupus familiaris) 

Mouse (Mus musculus) 

Rat (Rattus norvegicus) 

Rhesus monkey (Macaca mulatta) 

Zebrafish (Danio rerio) 

On Mouse Genome Informatics website on http://informatics.jax.org/, click on Download menu , click on Genes & Markers submenu, scroll down to Vertebrate Homology and then on HOM_AllOrganism.rpt file. When it opens in a browser, right click and save.

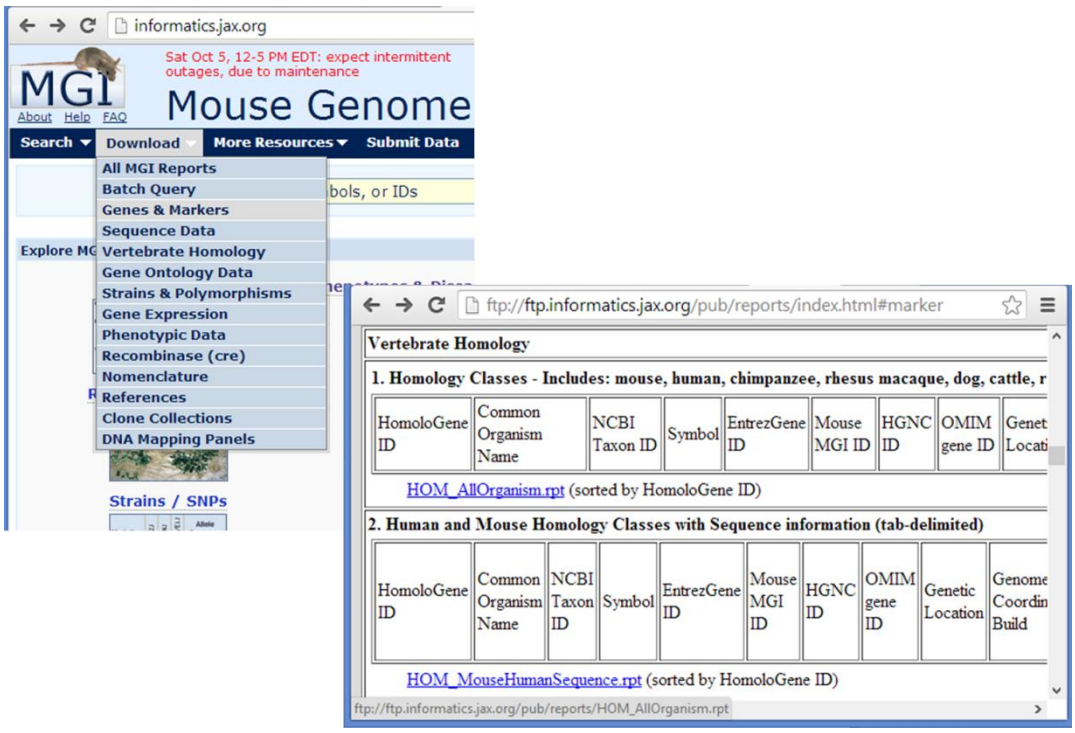


Figure 80: Accessing data from MGI website

Open the saved file in a spreadsheet (Microsoft Excel). I opened Excel application first before opening the text file. It would not work if done the other way round. Insert columns, add biological names, write formulae for the chromosomes and fill the missing rows as discussed under the Instruction manual section. Save the file as a text file ‘animals.txt’

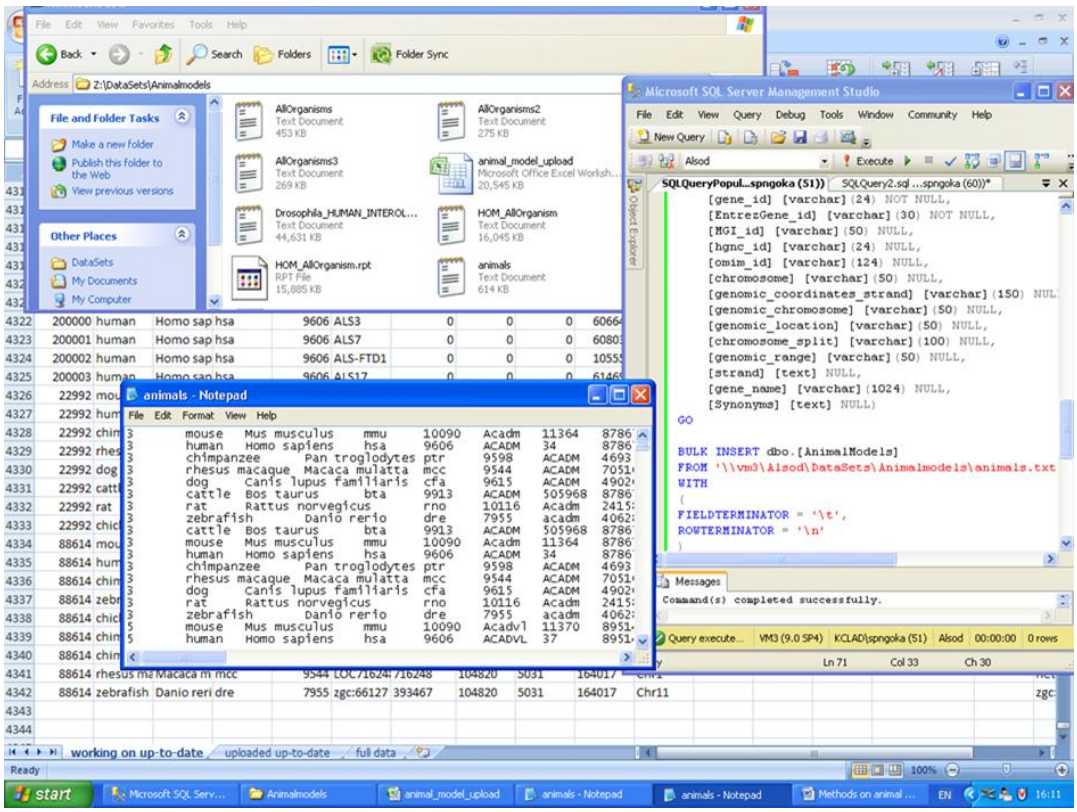


Figure 81: Screenshot on the process of creating Animal model table

Open a new query on database to create a table called ‘AnimalModels’, bulk insert the text file into the table as seen below:

```
CREATE TABLE [dbo].[AnimalModels](
    [HomoloGeneID] [varchar](50) NULL,
    [Organism_Name] [varchar](30) NOT NULL,
    [Biological_Name] [varchar](100) NOT NULL,
    [Kegg_id] [varchar](30) NULL,
    [ncbi_locuslink_id] [varchar](50) NULL,
    [gene_id] [varchar](24) NOT NULL,
    [EntrezGene_id] [varchar](30) NOT NULL,
    [MGI_id] [varchar](50) NULL,
    [hgnc_id] [varchar](24) NULL,
    [omim_id] [varchar](124) NULL,
    [chromosome] [varchar](50) NULL,
    [genomic_coordinates_strand] [varchar](150) NULL,
    [genomic_chromosome] [varchar](50) NULL,
    [genomic_location] [varchar](50) NULL,
```

```
	[chromosome_split] [varchar](100) NULL,
	[genomic_range] [varchar](50) NULL,
	[strand] [text] NULL,
	[gene_name] [varchar](1024) NULL,
	[Synonyms] [text] NULL)
GO

BULK INSERT dbo.[AnimalModels]
FROM '\\vm3\Alsod\DataSets\Animalmodels\animals.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.[AnimalModels]
ADD [ID] [int] IDENTITY(1,1) NOT NULL;
GO
```

6.3.10 Integrated Bioinformatics links

To avoid bias, users can retrieve broad information specific to the genes through external links which have been programmed automatically for each gene. Unique identifiers are utilized by systematically linking to broad databases and bioinformatics tools freely available online. Apart from links itemised below, ids were also derived from HGNC [525]. External links open in new windows and some of these bioinformatics tools are integrated into ALSoD using data from our database to create analysis. These links are categorised into three headings:

6.3.10.1 SCIENTIFIC

The scientific category is a list of hyperlinks associated with databases created by researchers for researchers.

Entrez Gene [526] – An NCBI information database centred on genes.

UCSC Browser [527] – Genome browser database by The University of California Santa Cruz for visualizing and querying data speedily at http://genome.ucsc.edu

Protein Structure[528] – Protein Domain and structure information as displayed on the UCSC Website

OMIM [529, 530] – It shows extensive information from the Online Mendelian Inheritance in Man database updated –though not regularly- by the research community.

Genecards[531] – These also show users an overview of gene information curated for research purpose.

Full Literature – This displays the result of a pubmed query e.g. for SOD1 gene, "SOD1[All Fields] AND amyotrophic lateral sclerosis/genetics[mh] AND motor neuron disease/genetics[mh] AND "humans"[MeSH Terms] AND english[la]"

ProtScale [532] – Links to an external website programmed to automatically use the uniprot identifier for the selected gene. This enables a computation and representation of the profile generated by an amino acid scale on the chosen protein.

KEGG [533] – shows the AA and NT sequence from the KEGG database to save valuable time spent by researchers in searching for sequences of specific genes online.

Uniprot [534] - The UniProt consortium provides access to the multiple data sets from the Swiss Institute of Bioinformatics (SIB), the European Bioinformatics Institute (EBI) and the Protein Information Resource (PIR).

iHop [535] – Information Hyperlinked over Proteins database is a vastly connected network making human literature research more intuitive and efficient using a text-mining method as described in their publication. This is linked from ALSoD to allow researchers an unbiased list of published materials relating to the selected gene.

Pathway in KEGG [533] – A KEGG pathway map of human Amyotrophic Lateral Sclerosis under the Neurodegenerative diseases of the Human disease section is hyperlinked from our website to display a pathway overview map.

GeneTest [536] - The GeneTests database and Web site which is hosted at NCBI displays gene reviews and information on laboratories and clinics related to the gene.

ALS review – A link to the bookshelf page of NCBI displaying regularly reviewed information on ALS overview to enable researchers view up-to-date research in the field of ALS.

AmiGO [537] – A link to this free open source web application developed and kept up-to-date by the GO Consortium lets users query, browse and view ontology and annotation data on definite genes.

Ensembl [538] – This displays a broad genome information system featuring an amalgamated set of genome annotation, databases, and other information.

NCBI [480] – This comprehensive database and website is linked from ALSoD showing a statistical overview of the dataset across its network of databases e.g. for ANG gene, the search query automatically coded to our database is “ANG HOMO ALS” displaying a total of 31 pubmed citations and abstracts.

Life Science DB (Japan) [485] – ALS mutation database developed in 2010 was constructed as part of the Life Science Integrated Database Project conducted by the Japan Ministry of Education, Culture, Sports, Science, and Technology was published. It contains their original experimental results and published data extracted from scientific journals. The database is expected to play a complementary role to the ALSoD database especially in collecting variations in the Asian region. A link to this website allows a wider perspective of mutations available in ALS without geographical bias. http://reseq.lifesciencedb.jp/resequence/GeneDetail.do?targetId=1&geneId=EG6647

ALSGene [539] – “ALSGene freely available on http://www.alsgene.org aims to serve as an exhaustive, unbiased, and regularly updated resource of genetic association studies in ALS. One of its key features will be up-to-date meta-analyses of all eligible genetic polymorphisms that have been investigated for association with ALS risk”.

Gen2Phen [557, 558] – Gen2Phen project is focussed on combining genetic variation databases of human and model organism by giving a complete view of Genotype-To-Phenotype (G2P) data.

Orphanet [559, 560] – This is an initiative in Europe aimed at improving the treatment and management of rare diseases. Information on rare diseases and orphan drugs are stored on a database www.orpha.net.

GeneNetwork [561] – This is an interactive tool which allows genetic network reconstruction without an proper knowledge of mathematical models.

MGI [562, 563] – Mouse Genome Informatics database resource contains information on mouse genes and markers.

LOVD [564, 565] – This is a Leiden Open-source Variation Database developed to store gene sequence variation information associated with human phenotype. It hosts 3,294 gene variant databases with 199,000 variants in 84,000 patients.

6.3.10.2 NON-SCIENTFIC

The non-scientific category is a list of hyperlinks associated with databases created by non-researchers for the public.

GeneWiki [540] – A hyperlinked website from ALSoD presents the free encyclopaedia of a selected gene from Wikipedia as this affords users the opportunity of viewing impartial details which may not be seen in scientific write-ups. E.g. http://en.wikipedia.org/wiki/SOD1

WolframAlpha [541] – With access to the freely available online computational knowledge engine, a link to the general knowledge about a gene is made available for proper understanding of a selected gene by non-scientists visiting the website. E.g. http://www.wolframalpha.com/input/?i=SOD1

WikiGenes [542] – A link to an active collaborative knowledge base for the life sciences has become an imminent scientific tool showing a references and broad information relating to psychiatry, chemical compound, biological context, anatomical context, associations, physical interactions, enzymatic interactions, regulatory relationships, high impact publications, analytical, diagnostic and therapeutic context of each gene. E.g. http://www.wikigenes.org/e/gene/e/6647.html

6.3.10.3 IN-HOUSE

The in-house category is a list of hyperlinks I designed within the database for easy access to details of data on ALSoD. These data are analysed and represented in various forms.

Mutations – Apart from a list of all genes available in a selected gene showing total number of mutations, mutation type, mutation location and a link to a brief detail of the mutations, a graphical overview of all the reported mutations for that gene and how they are dispersed along the codon are shown.

Mutation Google Map – A geographical overview of mutations in a selected gene are displayed on the globe in the form of labelled yellow pins representing countries where the mutations have been found. Users can interact with these pins to show the name of the country or click the pin to display all the mutations discovered

from the chosen country by splitting the pins into subs as seen in Figure 136. Due to software compatibility plug-in issues, this displays better in Internet Explorer.

Species – Users can view a multiple alignment of gene sequences. The Claustalw [566] which is a general purpose multiple sequence alignment program for DNA or proteins from EBI was used to provide multiple alignment for species on selected genes. The results from Claustalw are displayed colourfully in Jalview [551] which is a multiple alignment editor.

GeneMANIA – This was integrated into ALSoD to display predicted interactions between genes selected by checkboxes and passed into a the geneMANIA [509] web server framed and displayed on ALSoD website. It enables users to predict the interaction between genes by a few clicks on one page without going back and forth between websites. Due to technical difficulties, the analysis works better on Google chrome or Firefox web browsers than on Internet Explorer.

6.3.11 Borrowed Diagrams

Some images were designed either by colleagues or downloaded from the internet. These static images were redesigned with html codes to make them dynamic. A user can interact with the images with hyperlinks programmed on them which when clicked opens up detailed information.

6.3.11.1 Chromosomes

The image of each chromosome is imported from google image, downloaded to the PC and viewed on a powerpoint slide. Each gene location is mapped to the corresponding chromosome image with the gene name hyperlinked to its corresponding overview page on the database.

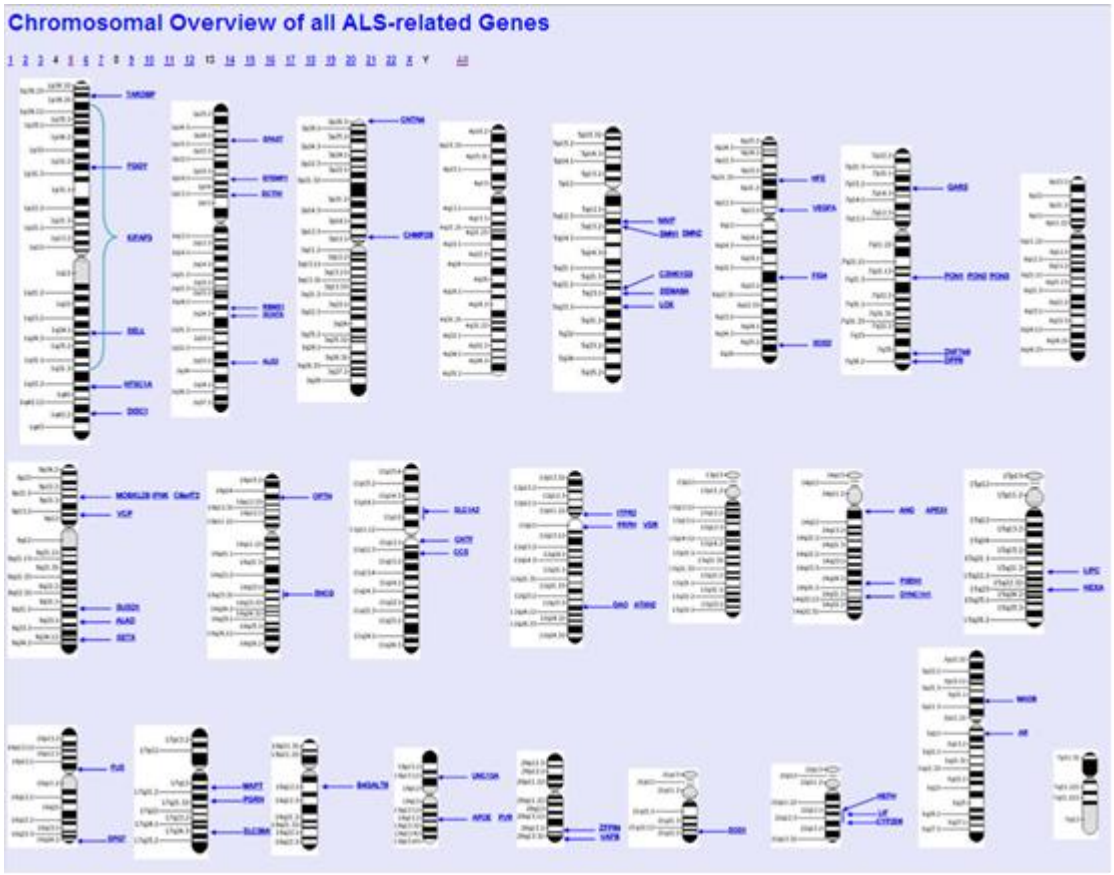


Figure 82: Chromosomal overview of all ALS-related genes

6.3.11.2 SOD1 subpsec score

The image was created to graphically display the location of the identified variants in SOD1 gene with its calculated subpsec score shown.

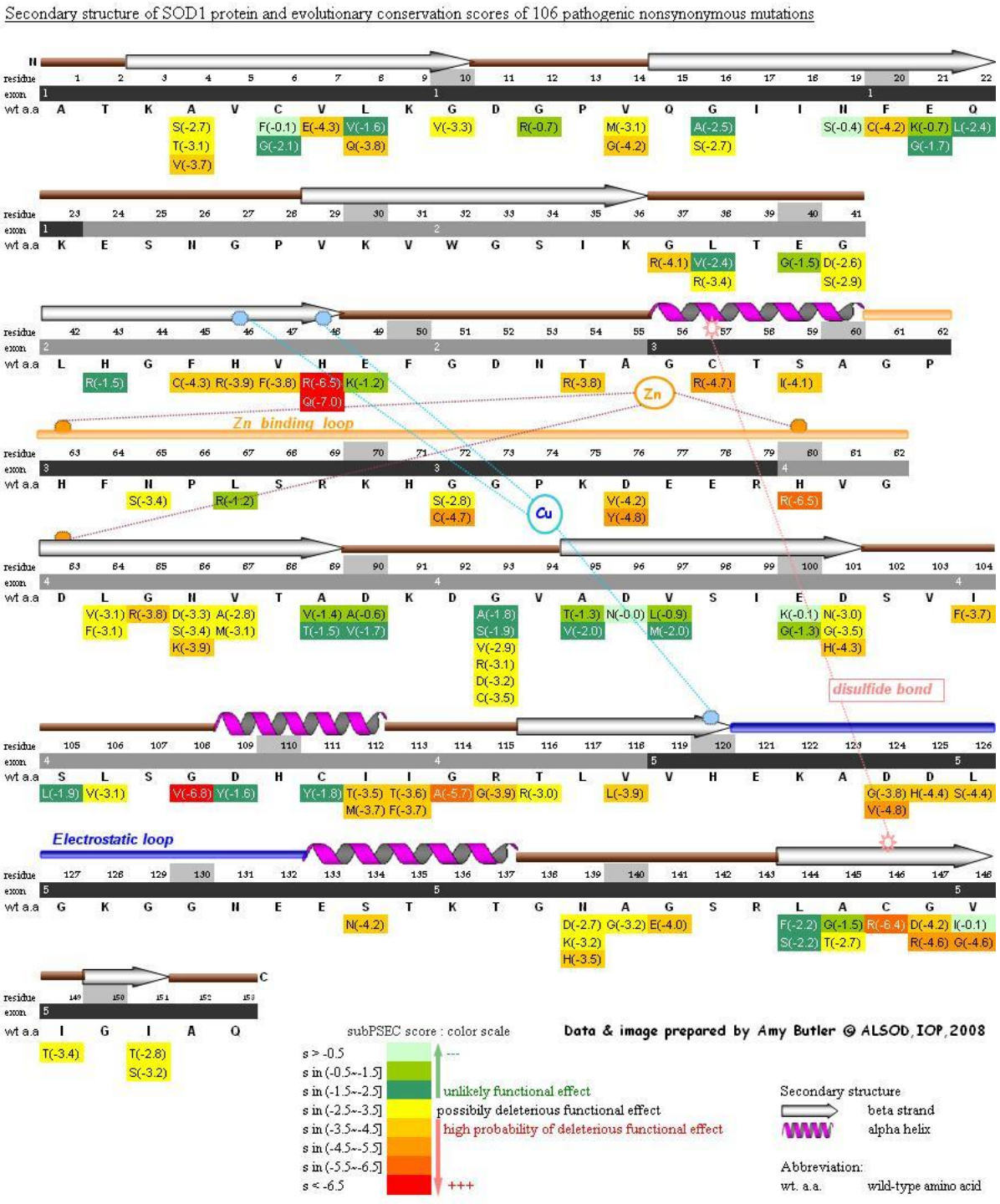


Figure 83: SOD1 subpsec score created in 2008 by Amy Butler

6.3.11.3 ALS-FTD relationship

This displays how ALS and FTD are related and compares the number of publications released on FTLD alone, ALS alone and ALS-FTD.

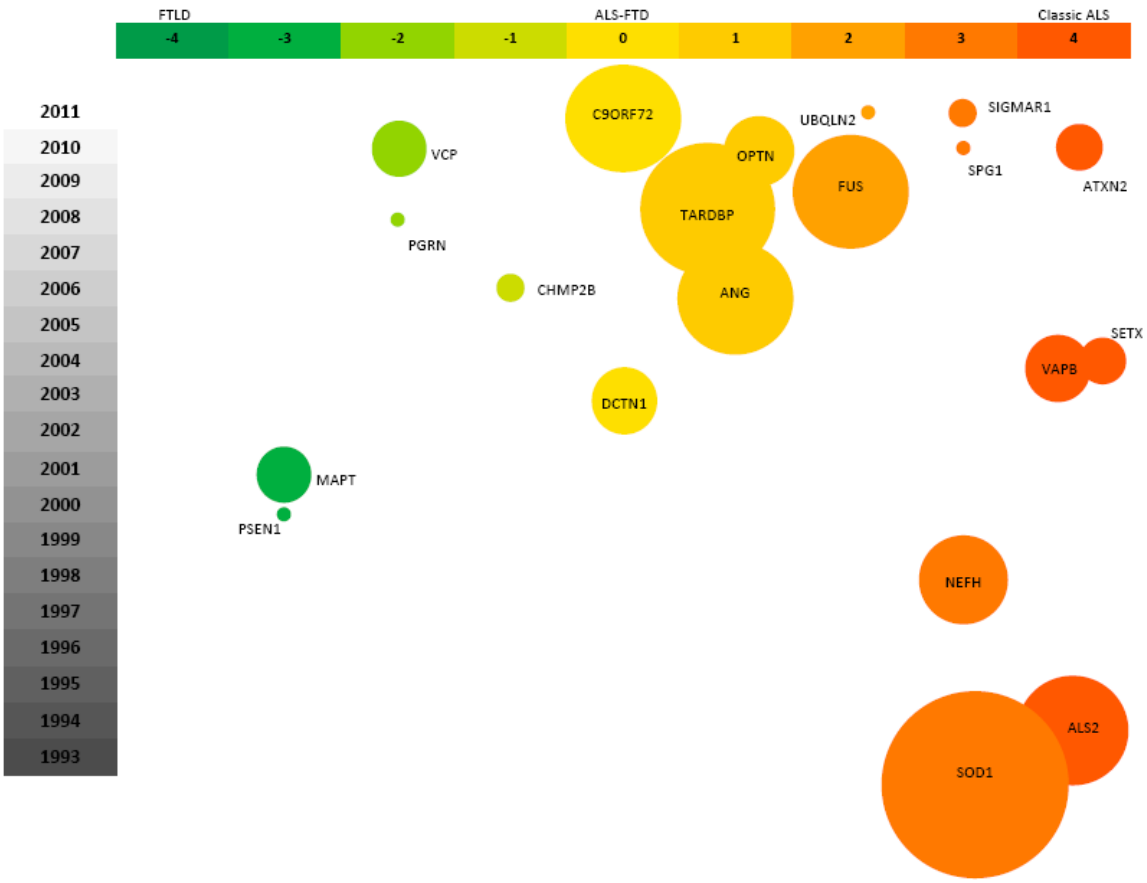


Figure 84: ALS-FTD relationship created in 2013 by Ashley Jones

6.3.12 Feedback

Feedback is gained from users of the database in four main ways:

6.3.12.1 Facebook

A Facebook page for ALSoD is found on the link http://www.facebook.com/srch.php#!/pages/ALSoD/307667685943735

```
<script>  (function(d, s, id) {
      var js, fjs = d.getElementsByTagName(s)[0];
      if (d.getElementById(id)) return;
      js = d.createElement(s); js.id = id;
      js.src = "//connect.facebook.net/en_GB/all.js#xfbml=1";
      fjs.parentNode.insertBefore(js, fjs);
  } (document, 'script', 'facebook-jssdk'));</script>

<div class="fb-like" data-href="http://alsod.iop.kcl.ac.uk/" data-send="true" data-layout="button_count" data-width="100" data-show-faces="false" data-action="recommend"></div>
```

6.3.12.2 Twiiter

```
<a class="twitter-timeline" href="https://twitter.com/ALSoD_Database" data-widget-id="382469573439549442">Tweets by @ALSoD_Database</a>

<script>!function(d,s,id){var js,fjs=d.getElementsByTagName(s)[0],p=/^http:/.test(d.location)?'http':'https';if(!d.getElementById(id)){js=d.createElement(s);js.id=id;js.src=p+"://platform.twitter.com/widgets.js";fjs.parentNode.insertBefore(js,fjs);}}(document,"script","twitter-wjs");</script>
```

6.3.12.3 Feedback page

This is a direct feedback page on the ALSoD website where forms are embedded. Comments are publicly displayed and a reCAPTCHA tool displays texts readable only by human users to prevent spammers from infiltrating the system. A news page generates automated summaries of ALS genetics news; and surveys

conducted through the freely available online survey tool ‘SurveyMonkey’ are embedded in the user interface as shown in Appendix 11.

6.3.12.4 Blog

A blog built in-house was developed to communicate with users on http://alsod.iop.kcl.ac.uk/BlogList.aspx. The script used to run the blog is shown in Appendix 12.

6.3.13 Tracking Visitors

6.3.13.1 Visitors’ statistics page

In addition, by tracking the registered country of origin of page viewing and download requests, accessibility of the ALSoD database to the international ALS community can be monitored directly. Codes have been described earlier under web design.

6.3.13.2 Google Analytics tool

ALSoD has a Google account and this enabled me to have access to the free-of-charge Google Analytics tool which is used to generate detailed statistics about a website's traffic. I logged in with my username and password, configured the step-by-step service in August 2012 and it keeps a daily log of activities taking place by users on ALSoD.

A code generated from the Google Analytics website was inserted on the MasterPage using:

```
<script type="text/javascript">
    var _gaq = _gaq || [];
    _gaq.push(['_setAccount', 'UA-34155904-1']);
    _gaq.push(['_trackPageview']);


    (function() {
        var ga = document.createElement('script'); ga.type = 'text/javascript'; ga.async = true;
        ga.src = ('https:' == document.location.protocol ? 'https://ssl' : 'http://www') + '.google-analytics.com/ga.js';
        var s = document.getElementsByTagName('script')[0]; s.parentNode.insertBefore(ga, s);
    })();
</script>
```

6.3.14 ChangeLog

6.3.14.1 ALSOD v0.1 Beta

First online in 1995 ALSOD (as formerly known) which was hosted at www.alsod.org was developed to store genetic and clinical information and to assist researchers in identifying correlations between phenotype and genotype in ALS for SOD1 mutations. The data available in the database were purely for the *SOD1* gene as this was the only available familial gene linked to ALS at the time [567].

6.3.14.2 ALSOD v1.0

In 1999, the database was first fully functional and available for the research community on www.alsod.org

6.3.14.3 ALSOD v2.0

In 2008, about 100 different mutation points across the *SOD1* sequence with corresponding clinical information were collated. Genetic mutations of the *SOD1* protein were linked to the hypothetical 3-D mutant structure hosted on a University College London server developed by Andrew Martin's team [568]. 50 users from 17 institutions registered with ALSOD to submit ALS patient and mutation data. 97 familial individuals with 122 mutation data on *SOD1* were stored.

The website was relocated to http://alsod.iop.kcl.ac.uk/als in 2007 following the loss of alsod.org domain. Data could be downloaded freely and the database queried to look for a specific mutation type in four ALS genes (*SOD1, ALS2, VAPB, NEFH*) or for specific information on patient data.

6.3.14.4 ALSOD v2.5

In 2010, the database had records for 52 genes, 13 genome-wide association studies, 254 mutation data and 255 patient data. This required continued curation and maintenance to remain up-to-date. While I have automated much of this, there was still a need for human review, selection and intervention.

6.3.14.5 ALSoD v 3.0

ALSoD (ALS online genetics Database) is now a relational database with a massive increase in available data through submissions by researchers and regular update by the database curators. The schema has been redesigned for uncomplicated future addition of familial and sporadic ALS patient data, associated mutations and published ALS genes.

6.3.14.6 ALSoD v 4.0 current structure

The ALS Online genetics Database (formerly hosted on www.alsod.org) is currently available on http://alsod.iop.kcl.ac.uk providing detailed comparative analysis, gene overview, literature curation and

integrated bioinformatics tools. The main entry point to the database is a two-section list of all ALS-related genes leading to the gene overview page providing a uniform display of information about a gene.

6.3.15 Search keyword(s) externally or internally

On the masterpage, two search buttons are embedded to allow users search ALS keywords on other databases and/or within ALSoD. The first search button searches the keyword(s) typed into the textbox next to it in the Google search engine to produce results embedded on the webpage for viewing. The second search box opens a new page and searches the webpages of ALSoD website for the keyword(s) using Google plugin customized to search ALSoD. The result is displayed on a floating page.

Figure 85: Search externally and internally

6.3.15.1 Externally

```
<td align="right"><span lang="en-gb">
        <asp:Button runat="server" id="sa" text="Search" OnClick="SearchSite" style="margin-top: 5px" Height="21px" Width="57px" />
        <asp:TextBox runat="server" id="q" style="margin-left: 0px; margin-right: 1px;" />
</td>
```

6.3.15.2 Internally

```
<td align="right"><span lang="en-gb">
    <asp:Button runat="server" id="sa2" text="Search" OnClick="SearchSite2" style="margin-top: 5px" Height="21px" Width="57px" />
   <asp:TextBox runat="server" id="q2" style="margin-left: 0px; margin-right: 1px;" />
</td>
```

6.4 Credibility analysis of putative disease-causing genes using bioinformatics

Overview of the method is graphically shown below in Figure 86.

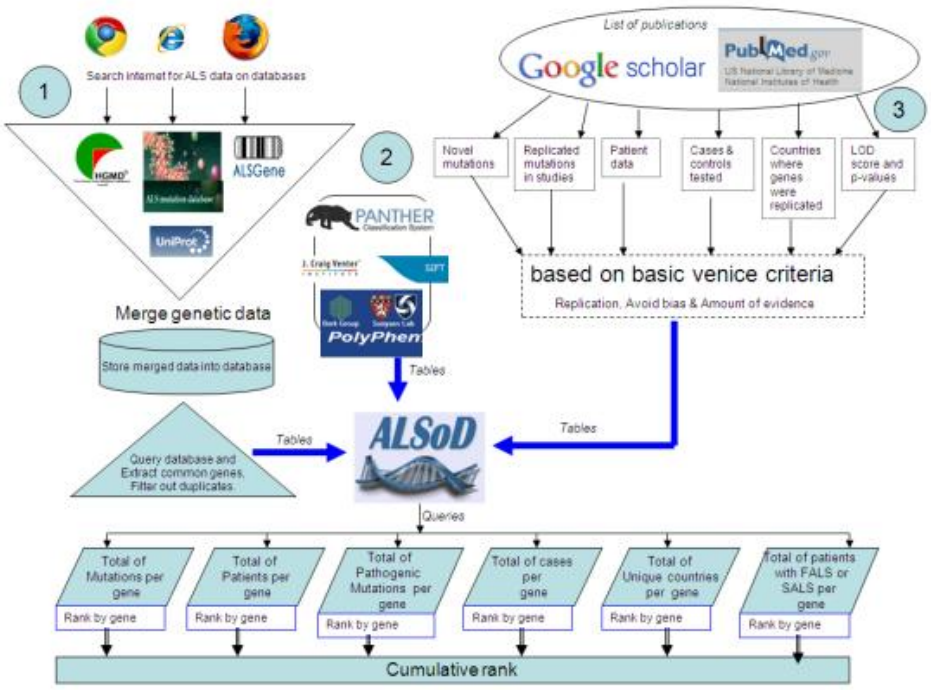


Figure 86: Overview of credibility analysis method

6.4.1 Data collection

Genes with at least one publication suggesting involvement in adult onset familial ALS were studied [108]. I excluded genes with limited clinical data, absent mutational data or unreplicated results. Publicly listed variants for the included genes derived from ALSGene, Uniprot, ALS Mutation and HGMD databases were merged with variant lists in ALSoD, and filtered for duplicates (Figure 1).

6.4.2 Pathogenicity analysis using bioinformatic tools

PANTHER (Protein Analysis Through Evolutionary Relationships) [548], SIFT (Sorting Intolerant From Tolerant) [549] and POLYPHEN (Polymorphism Phenotyping) [569] programs were used to analyse variants for possible pathogenicity. These tools generated a set of scores for the variants analysed, which for PANTHER are given as a subSPEC (substitution position-specific evolutionary conservation) score and for POLYPHEN given as score differences for PSIC (position-specific independent counts). In PANTHER, all possible mutations for each gene were generated using perl scripts and run on the web service in batches. SubPSEC scores <=-5.0 were defined as damaging and subPSEC scores >-5.0 defined as not

damaging. In SIFT, all possible unique codons in each gene were generated using perl script with scores <=0.05 defined as damaging and scores >0.05 defined as not damaging. In POLYPHEN, all mutations available in a gene on ALSoD were run through the web service one after the other and PSIC score differences >=1.5 defined as damaging and PSIC score differences <1.5 as not damaging.

6.4.2.1 Write scripts to generate all possible mutations

Using tdp43 gene as the sample gene, access the overview page of the gene using http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=TARDBP(TDP43). Under bioinformatics section, select sequence link which opens the KEGG[533] page. The process of running and combining scripts to generate the end result is as shown in Appendix 33

.

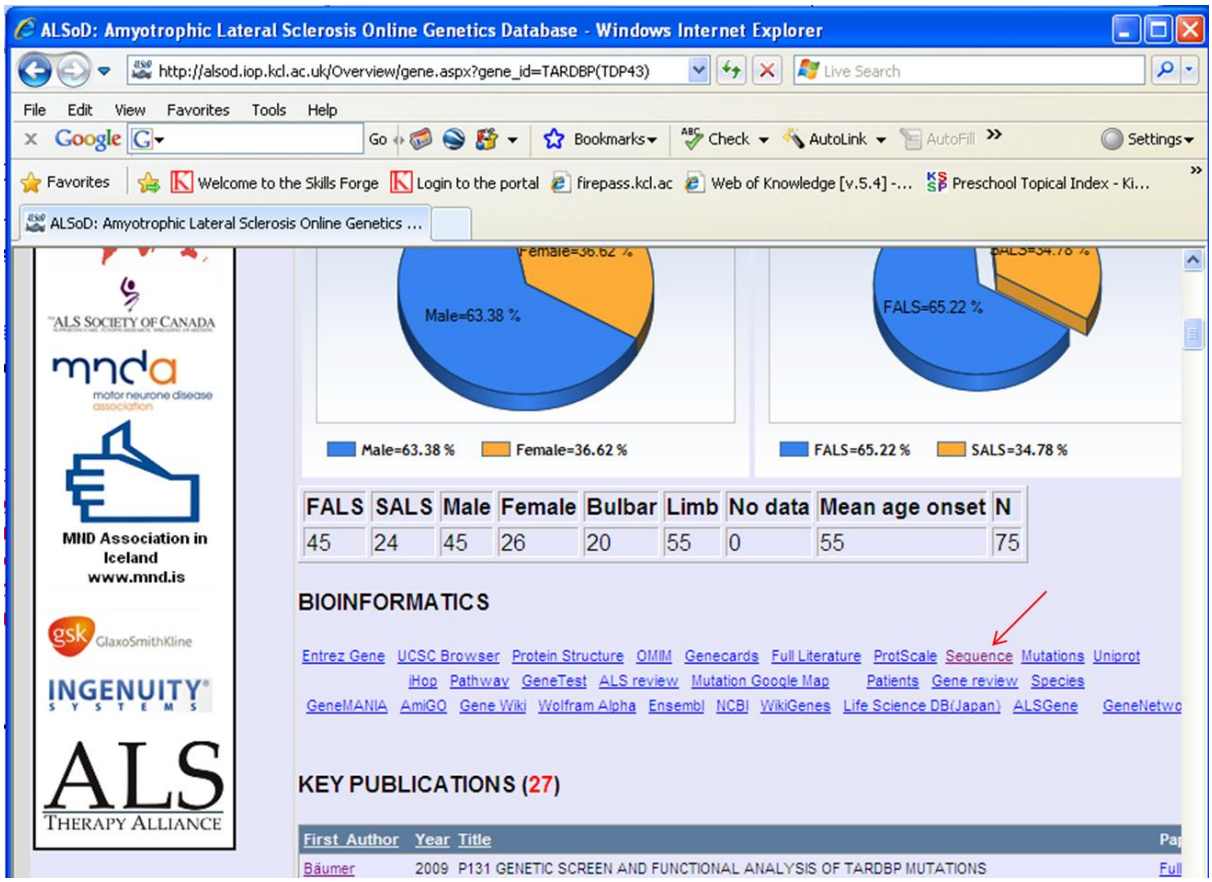


Figure 87: Generating all possible gene mutation process 1

From the KEGG page, select, copy and paste the Amino Acid sequence to an empty text file and save as tdp_codon.txt

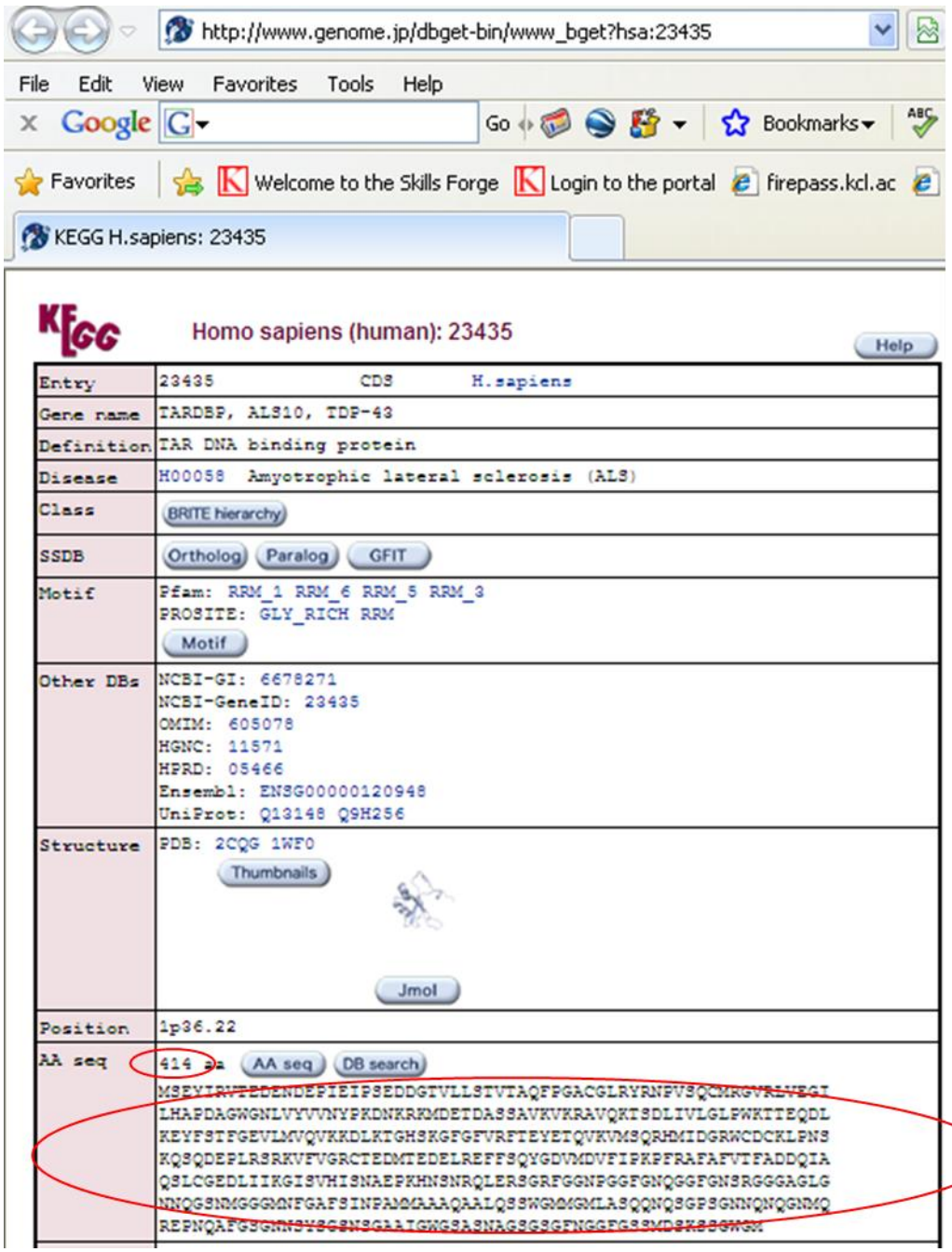


Figure 88: Generating all possible gene mutation process 2

Once the TDP43_codon.txt is saved, open the perl script for transposing the sequence vertically named as transpose_sequence.pl

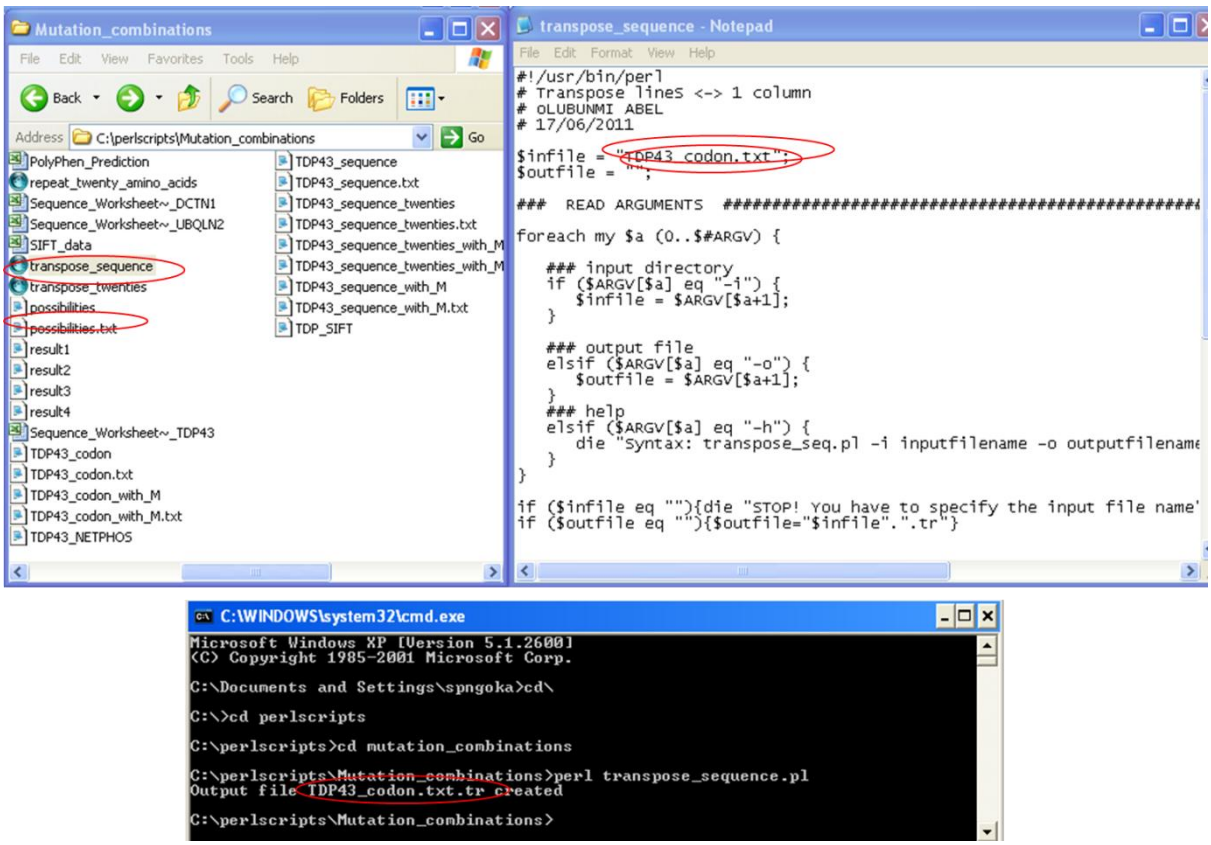


Figure 89: Generating all possible gene mutation process 3

The script generates a file TDP43_codon.txt.tr displaying a single column. This single column of data was copied from the text file to a new excel spreadsheet. This column was replicated across 20 columns as shown below in Figure 90.

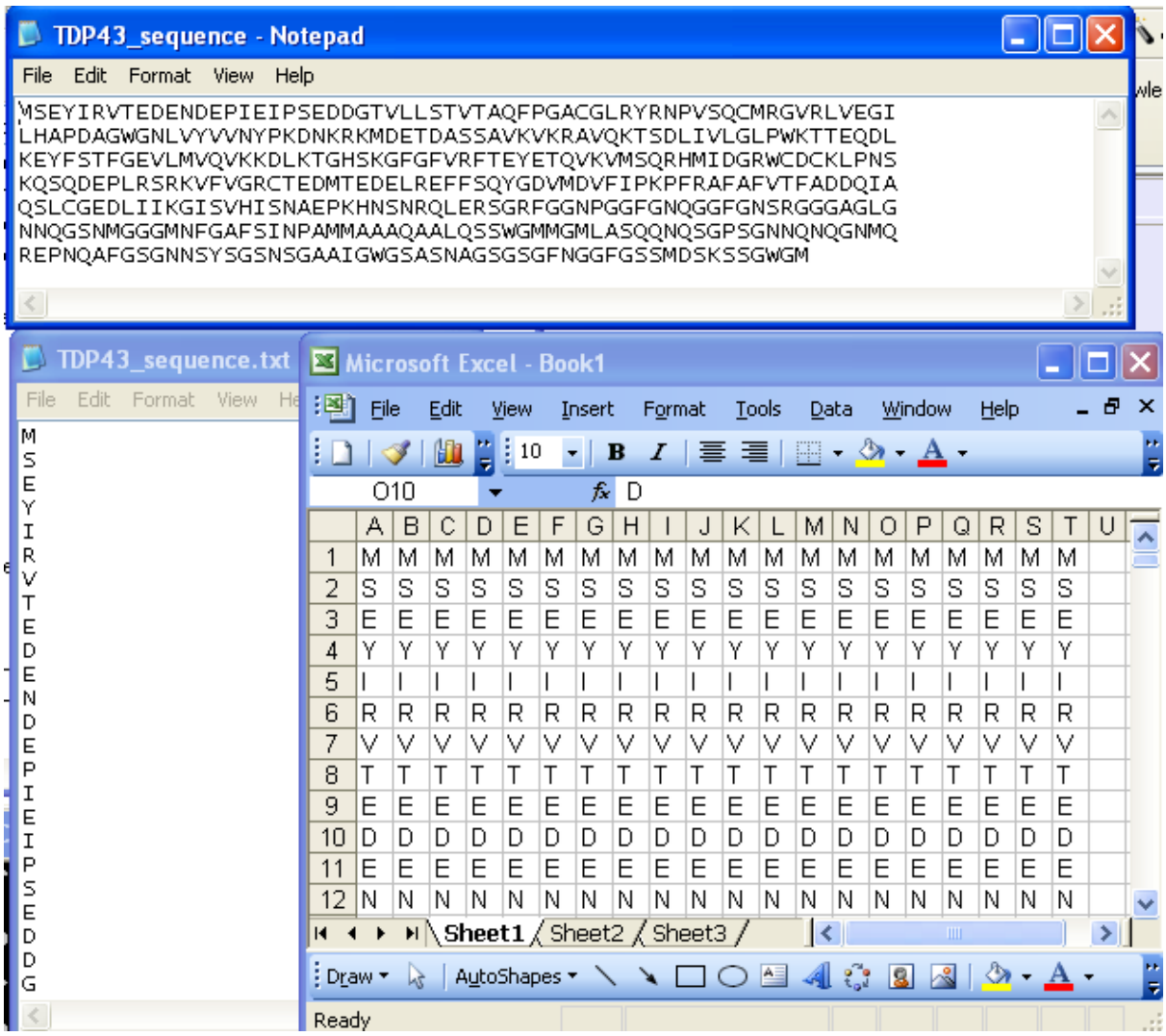


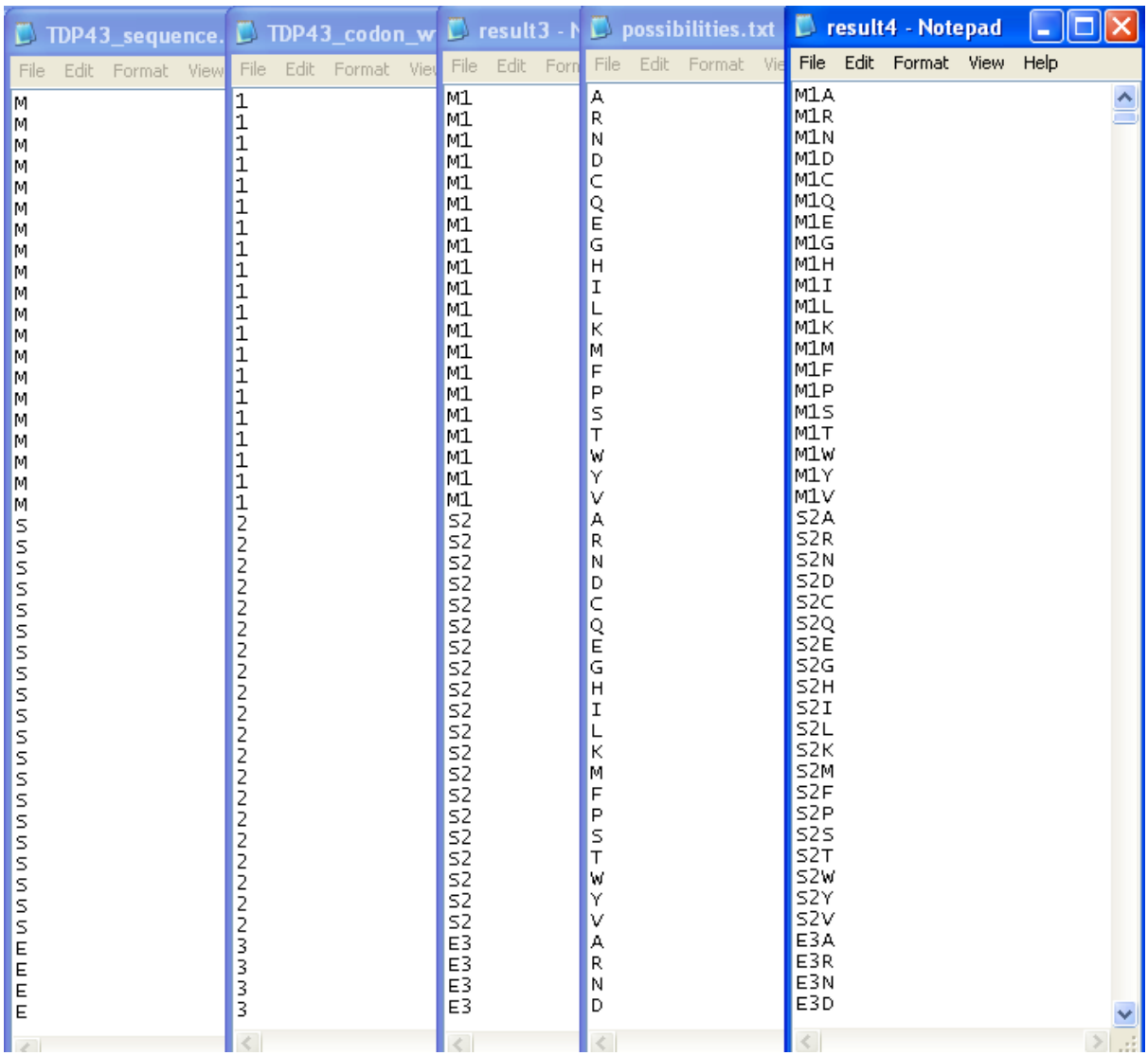
Figure 90: Generating all possible gene mutation process 4

To create 153 of the 20 amino acids ARNDCQEGHILKMFPSTWYV, run script repeat_twenty_amino_acids.pl to produce possibilities.txt, then run transpose_twenties.pl with infile as possibilities to generate possibilities.txt.tr

To combine the three files generated, run join_column_notabs.pl script :

TDP43_sequence_twenties.txt.tr + TDP43_codon.txt.tr = result3.txt

result3.txt + possiblilities.txt.tr = result4.txt


TDP43_sequence.
File Edit Format View
M M M

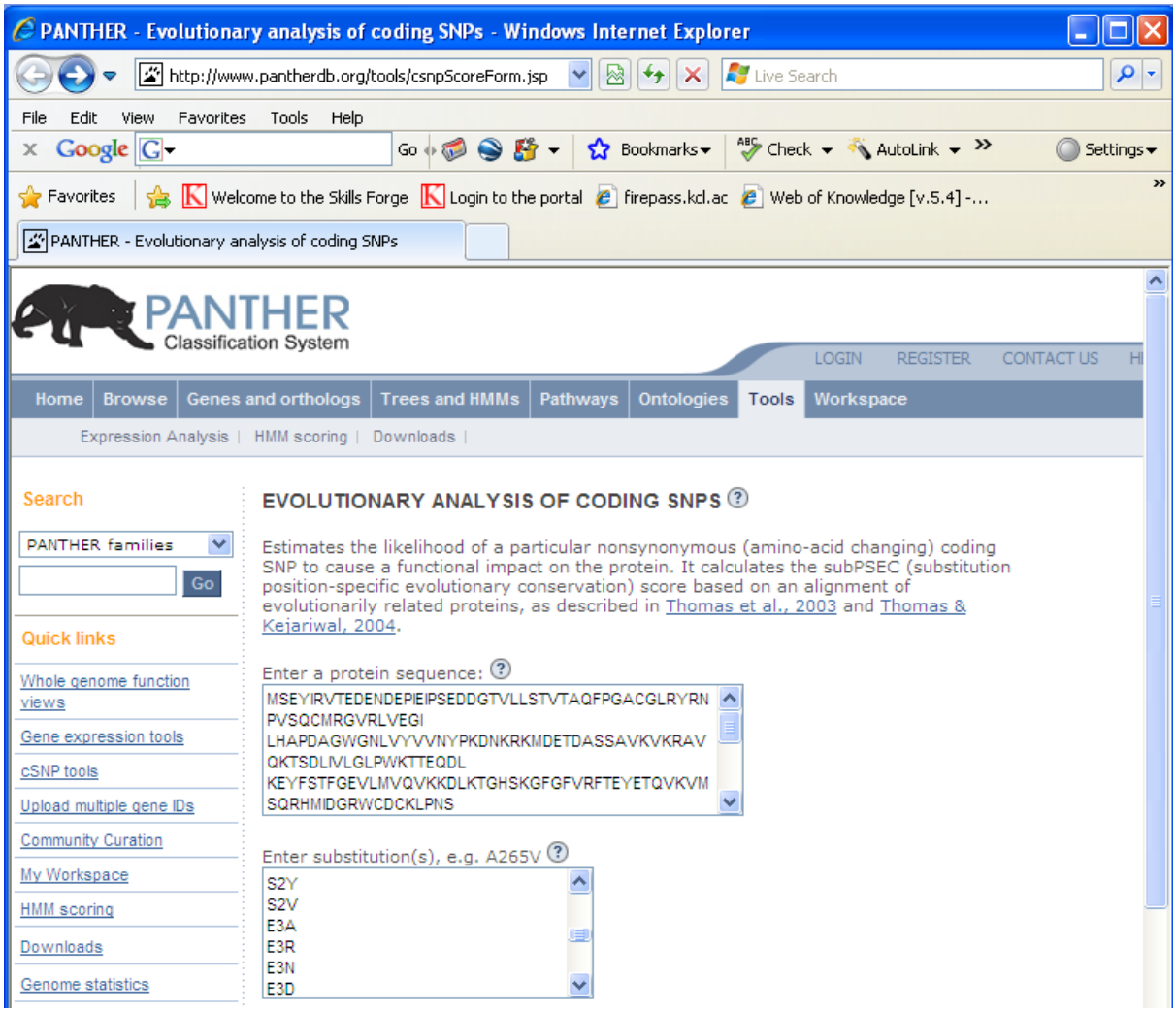
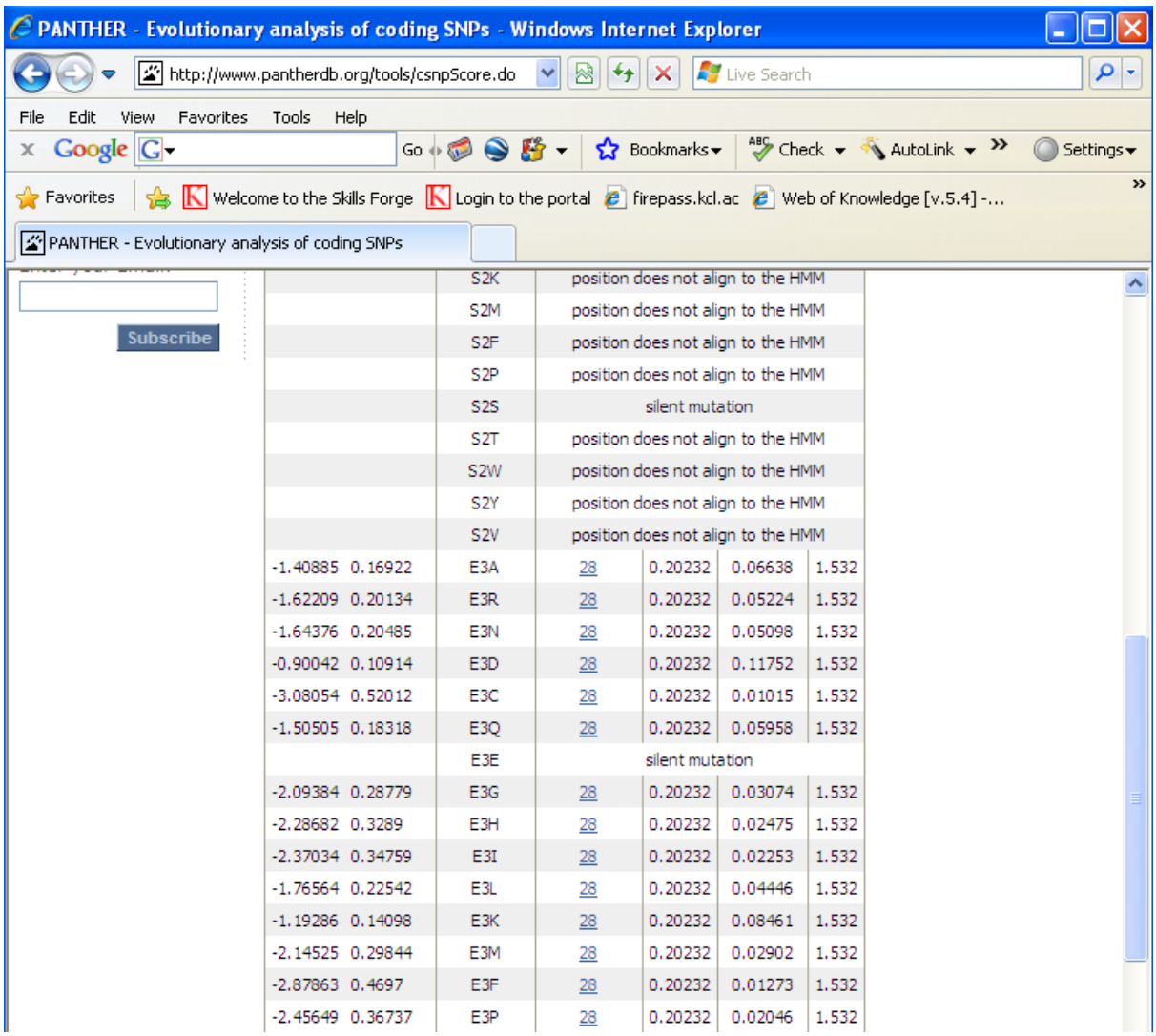


Figure 92: Pathogenicity prediction in PANTHER process 1

Wait for approximately 3mins to see the batch of analysis generated. Open the generated file known as



	S2K	position does not align to the HMM				
	S2M	position does not align to the HMM				
	S2F	position does not align to the HMM				
	S2P	position does not align to the HMM				
	S2S	silent mutation				
	S2T	position does not align to the HMM				
	S2W	position does not align to the HMM				
	S2Y	position does not align to the HMM				
	S2V	position does not align to the HMM				
-1.40885	0.16922	E3A	28	0.20232	0.06638	1.532
-1.62209	0.20134	E3R	28	0.20232	0.05224	1.532
-1.64376	0.20485	E3N	28	0.20232	0.05098	1.532
-0.90042	0.10914	E3D	28	0.20232	0.11752	1.532
-3.08054	0.52012	E3C	28	0.20232	0.01015	1.532
-1.50505	0.18318	E3Q	28	0.20232	0.05958	1.532
		E3E	silent mutation			
-2.09384	0.28779	E3G	28	0.20232	0.03074	1.532
-2.28682	0.3289	E3H	28	0.20232	0.02475	1.532
-2.37034	0.34759	E3I	28	0.20232	0.02253	1.532
-1.76564	0.22542	E3L	28	0.20232	0.04446	1.532
-1.19286	0.14098	E3K	28	0.20232	0.08461	1.532
-2.14525	0.29844	E3M	28	0.20232	0.02902	1.532
-2.87863	0.4697	E3F	28	0.20232	0.01273	1.532
-2.45649	0.36737	E3P	28	0.20232	0.02046	1.532

Figure 93: Pathogenicity prediction PANTHER process 2

Select the next 1000 lines of possible substitutions, once result is generated, select all, copy and paste on a new line at the end of the pantherCsnpAnalysis.txt file

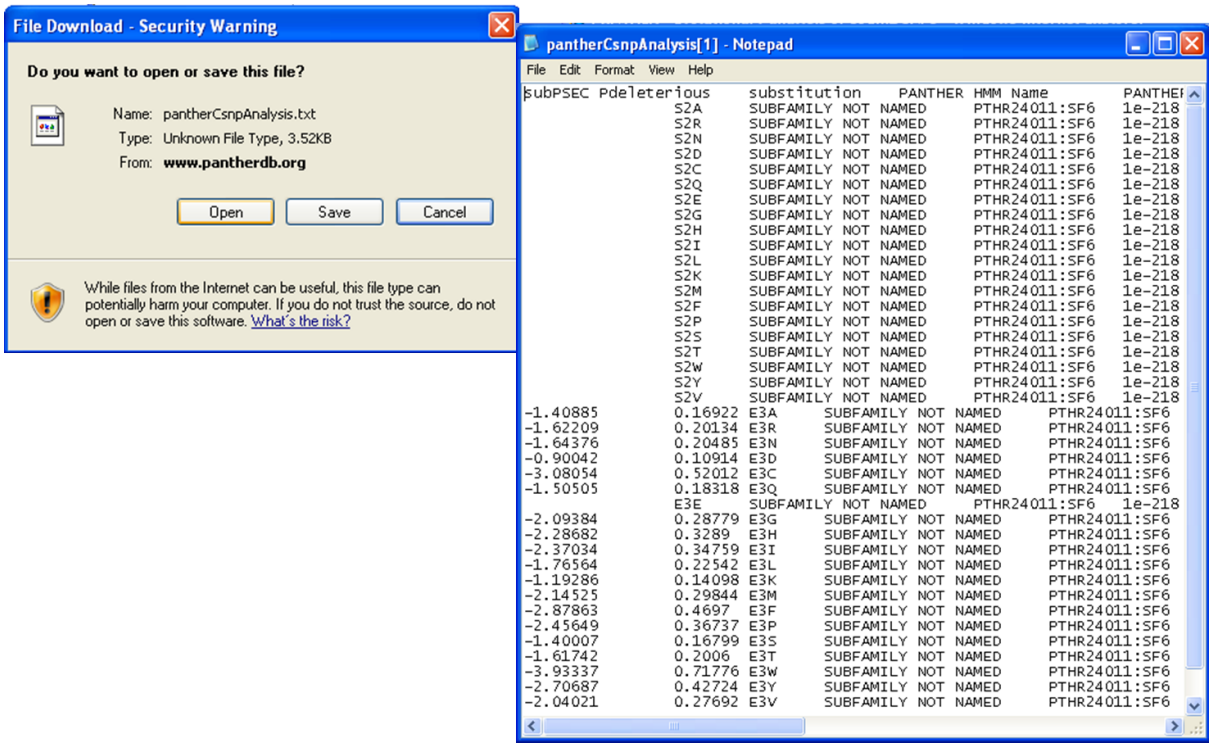


Figure 94: Pathogenicity prediction in PANTHER process 3

Repeat last step until all possible substitutions have been submitted for analysis.

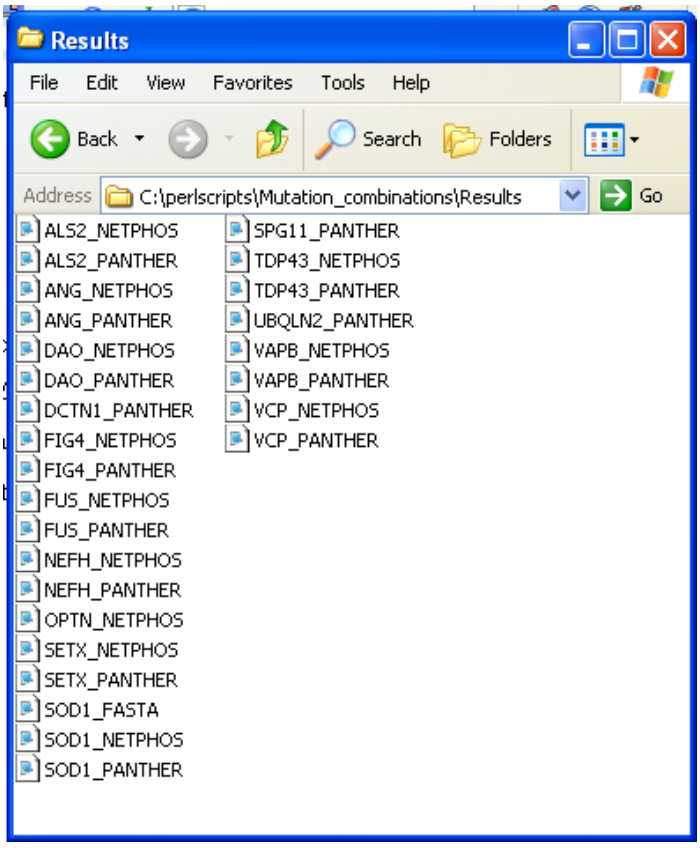


Figure 95: Pathogenicity prediction in PANTHER process 4

6.4.2.3 SIFT results submitted and data extracted

On a browser, open the url to SIFT (*http://sift.jcvi.org/www/SIFT_BLink_submit.html*) and under the 'Protein Sequence ID' section, type the NCBI Gi_number for TDP43 which is 49065494 (acquired from the NCBI website). Also, browse the filename 'result4.txt' and submit for analysis.

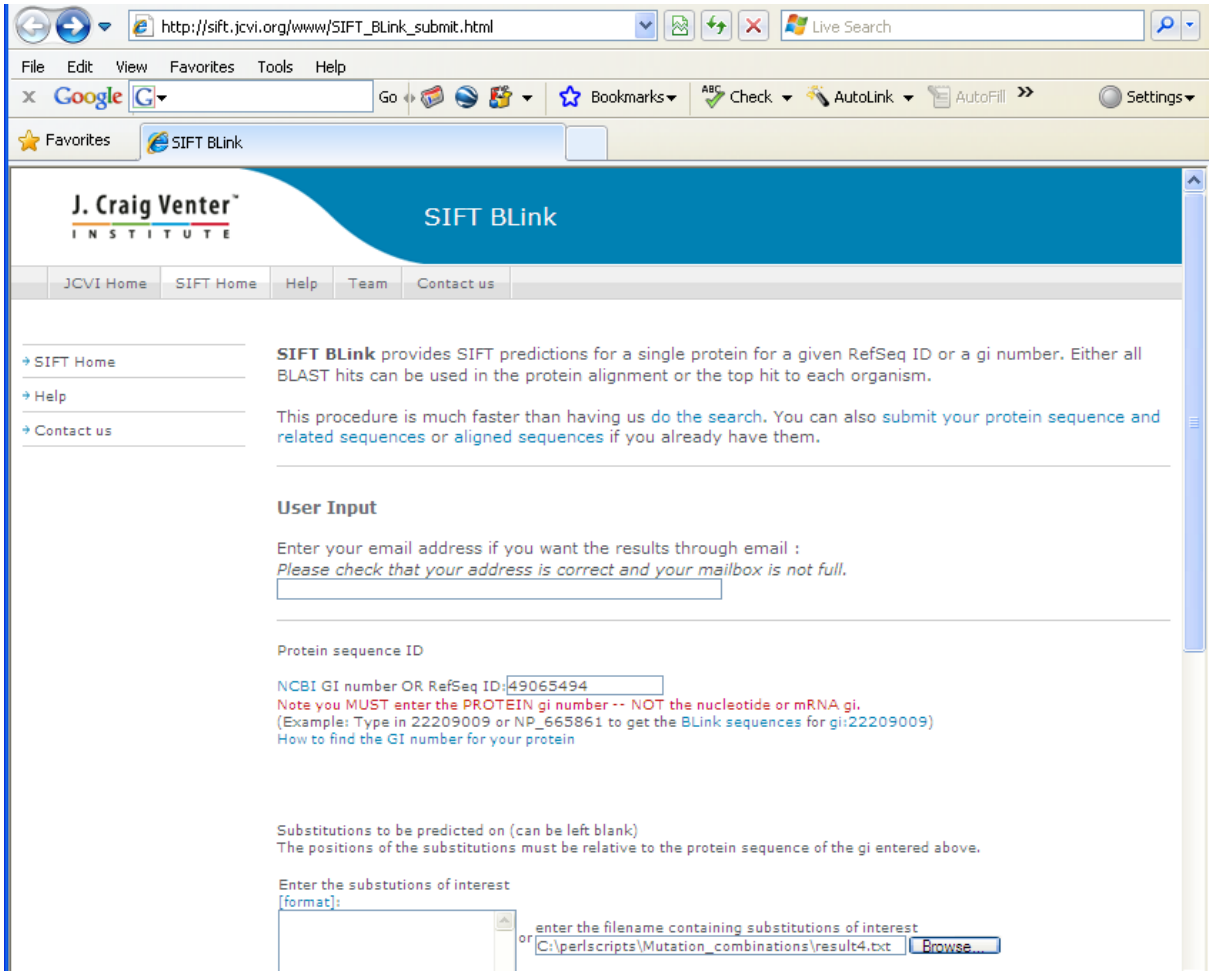


Figure 96: Pathogenicity prediction in SIFT process 1

Wait for about 2 mins and the result is generated.

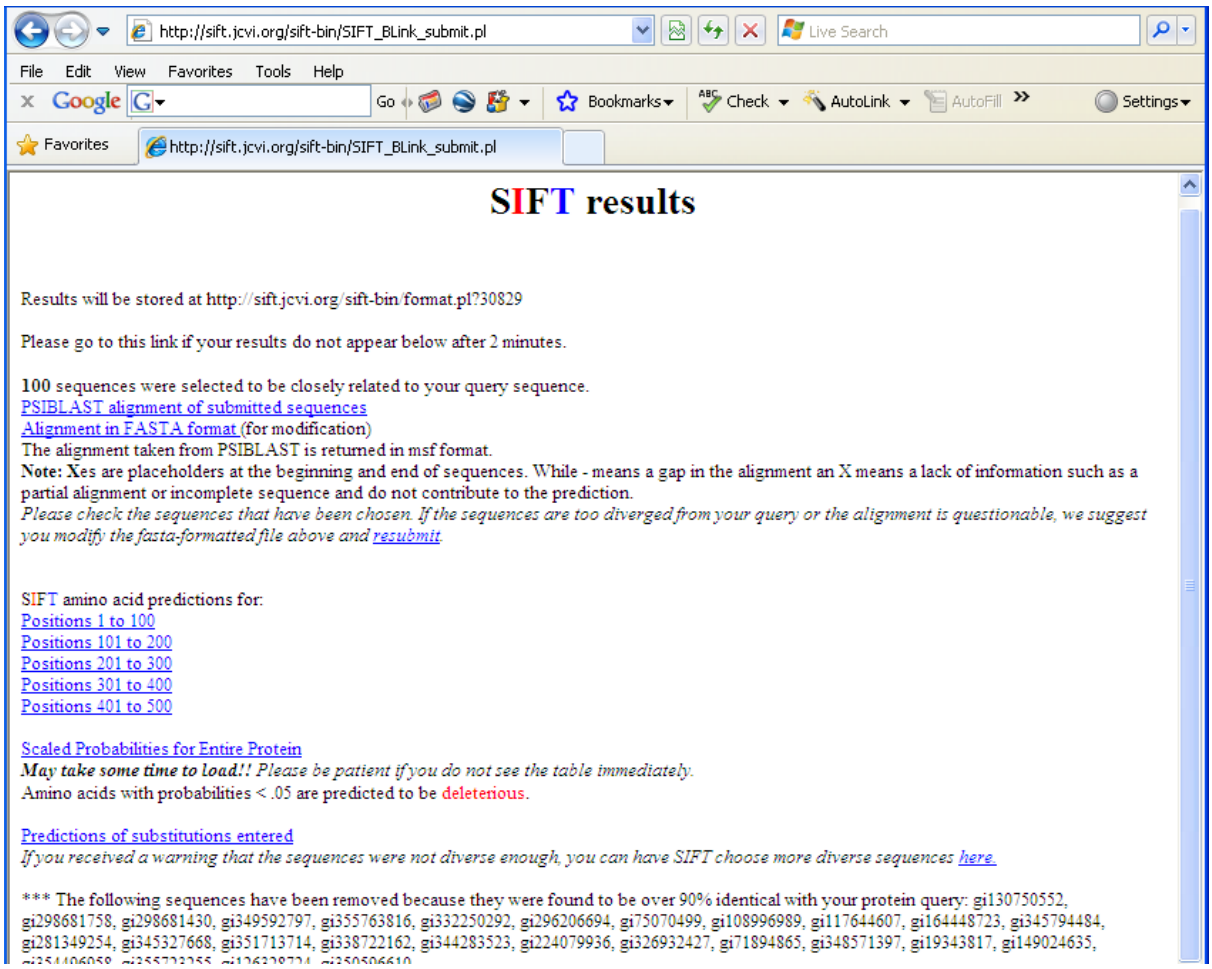


Figure 97: Pathogenicity prediction in SIFT process 2

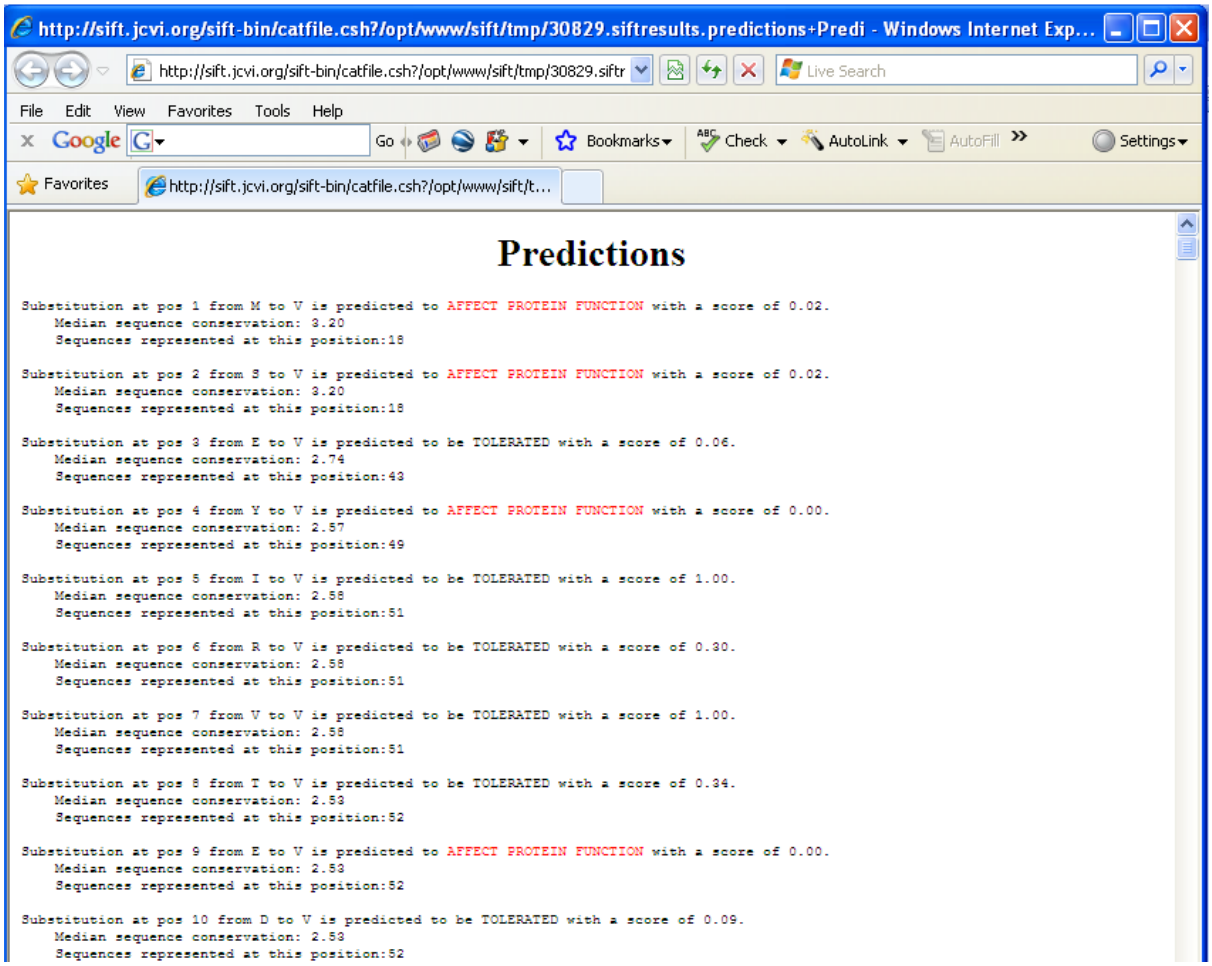
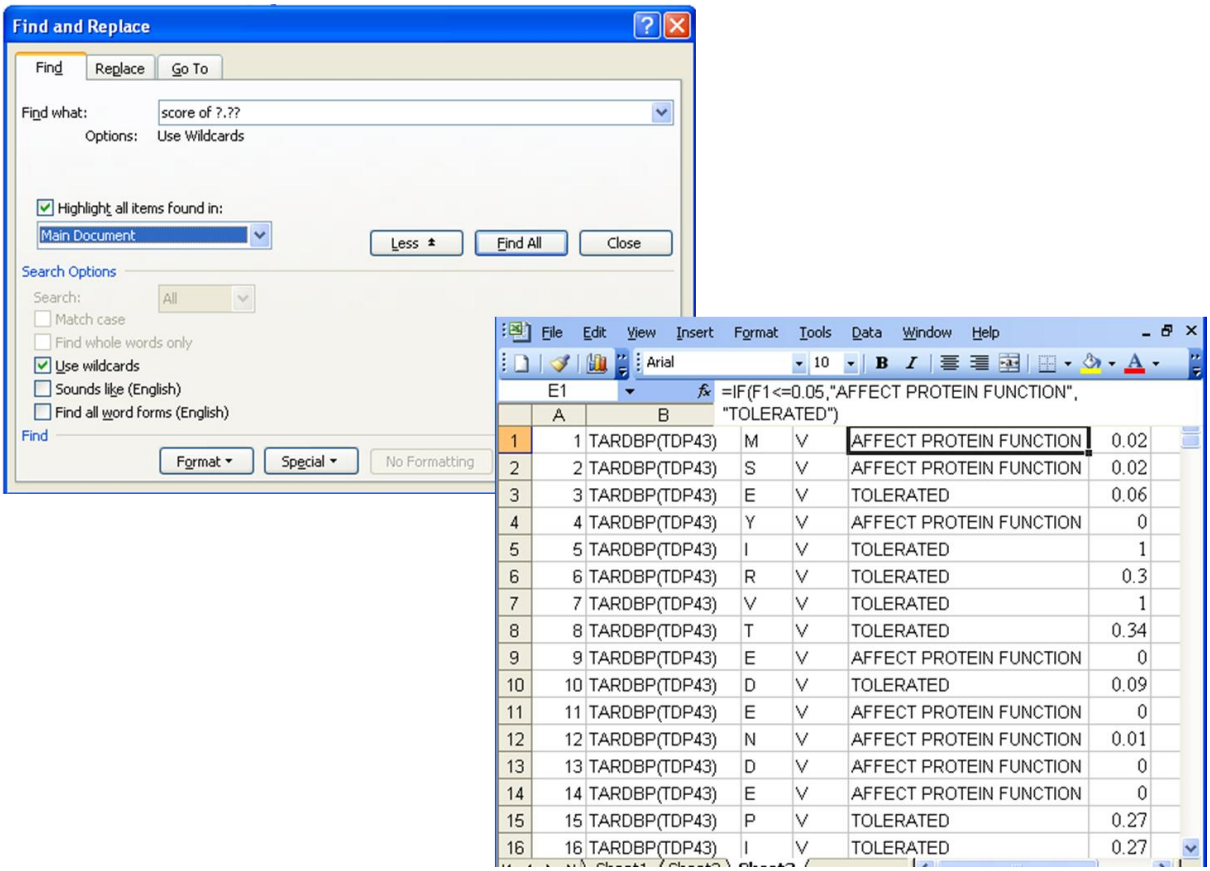


Figure 98: Pathogenicity prediction in SIFT process 3

To prepare the generated information for a database-acceptable format, select all, copy and paste in Microsoft word. Under the home or edit menu (depending on the version of MS Word used or CTRL H) which pos up the Replace dialogue box, under the find what section, type ***score of ?.??*** , check the highlight all items found, check use wildcards. This selects all the scores generated , copy and paste in a new spreadsheet. Add a column of TDP43's amino acids, a column for V replicated to the end of the row. Another row of function describing the result of each row is created using the following excel function:

IF (F1<=005, "AFFECT PROTEIN FUNCTION", "TOLERATED")



	A	B				
1	1	TARDBP(TDP43)	M	V	AFFECT PROTEIN FUNCTION	0.02
2	2	TARDBP(TDP43)	S	V	AFFECT PROTEIN FUNCTION	0.02
3	3	TARDBP(TDP43)	E	V	TOLERATED	0.06
4	4	TARDBP(TDP43)	Y	V	AFFECT PROTEIN FUNCTION	0
5	5	TARDBP(TDP43)	I	V	TOLERATED	1
6	6	TARDBP(TDP43)	R	V	TOLERATED	0.3
7	7	TARDBP(TDP43)	V	V	TOLERATED	1
8	8	TARDBP(TDP43)	T	V	TOLERATED	0.34
9	9	TARDBP(TDP43)	E	V	AFFECT PROTEIN FUNCTION	0
10	10	TARDBP(TDP43)	D	V	TOLERATED	0.09
11	11	TARDBP(TDP43)	E	V	AFFECT PROTEIN FUNCTION	0
12	12	TARDBP(TDP43)	N	V	AFFECT PROTEIN FUNCTION	0.01
13	13	TARDBP(TDP43)	D	V	AFFECT PROTEIN FUNCTION	0
14	14	TARDBP(TDP43)	E	V	AFFECT PROTEIN FUNCTION	0
15	15	TARDBP(TDP43)	P	V	TOLERATED	0.27
16	16	TARDBP(TDP43)	I	V	TOLERATED	0.27

Figure 99: Pathogenicity prediction in SIFT process 4

The final spreadsheet is copied and pasted into a text file named ‘SIFT_Prediction_TDP43.txt’

6.4.2.4 PolyPhen results submitted one by one for existing mutations

On a browser, open the url to PolyPhen (*http://genetics.bwh.harvard.edu/pph/*) and under Query Data section e.g for A4V in SOD1

Protein or SNP identifier : P00441

Position: 5

Substitution: AA1 is A and AA2 is V

Query description: A4V

Submit Query button

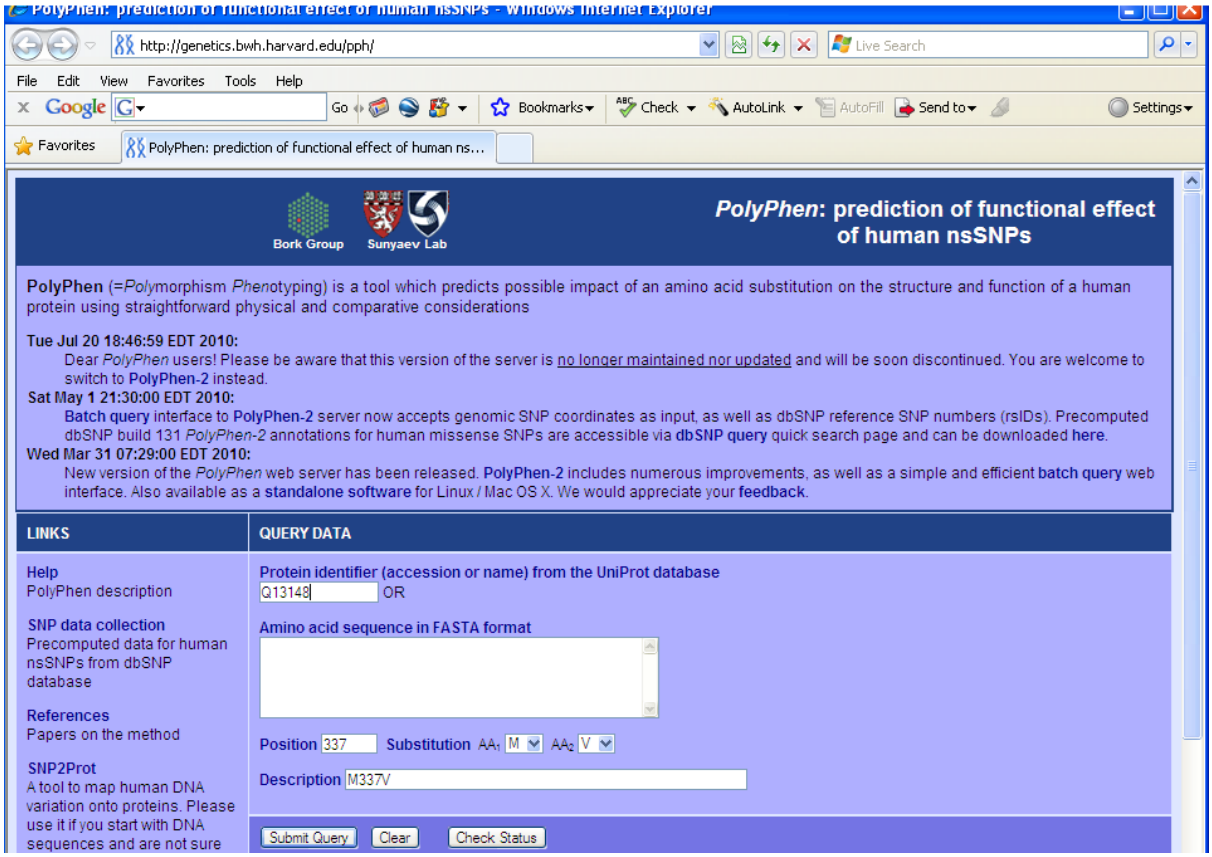


Figure 100: Pathogenicity prediction in PolyPhen process 1

The PolyPhen server was later upgraded to PolyPhen-2 [554] version 2.2.2 on 15th February 2012 to with a new url (http://genetics.bwh.harvard.edu/pph2/index.shtml). Processing a variant for prediction is the same but the scores are sometimes different. The “HumDiv” prediction was used rather than the “HumVar”

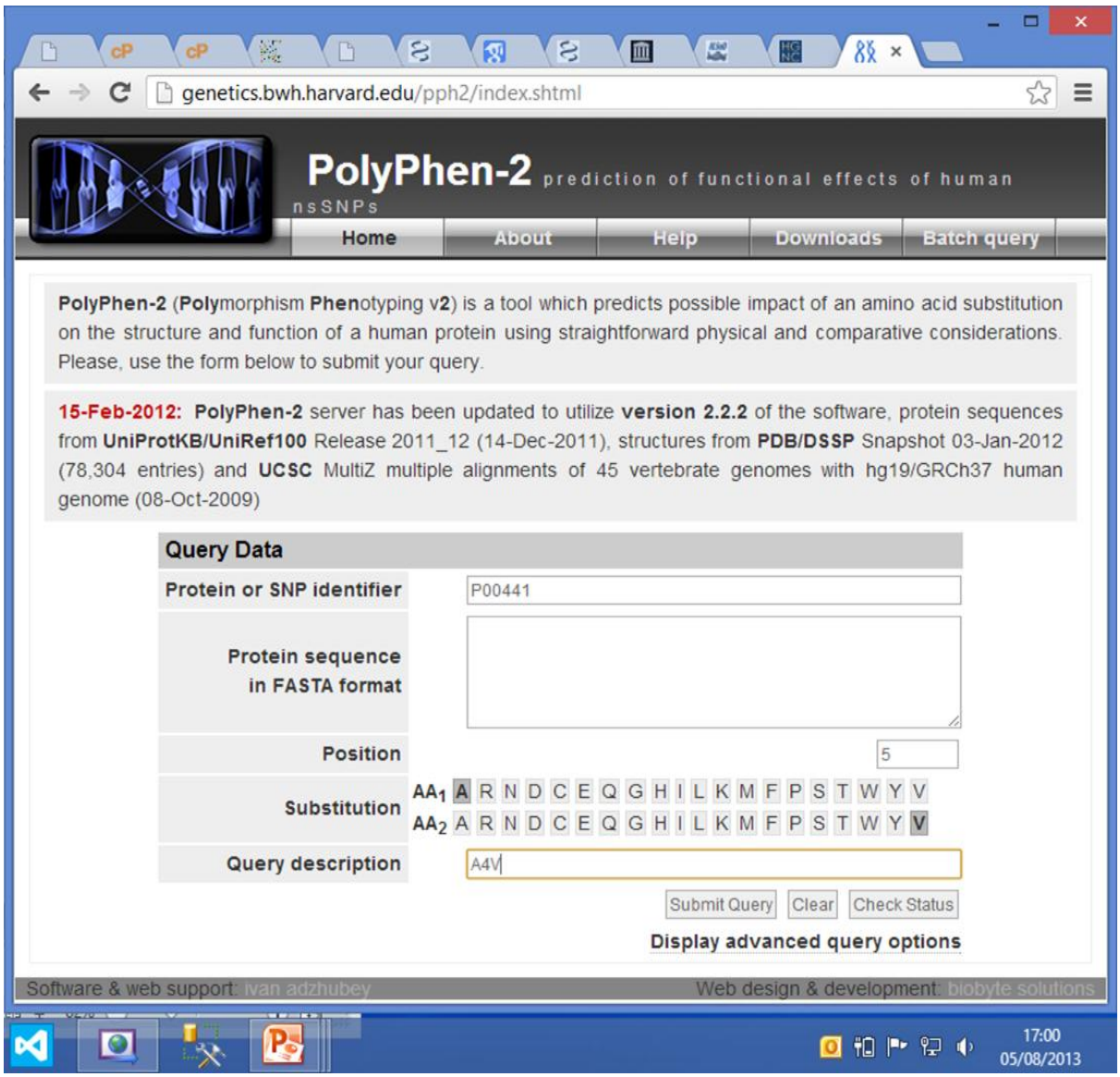


Figure 101: Pathogenicity prediction in PolyPhen process 2

The Gateway interface displaying the status of request sent to the server is shown. If an error is discovered, it shows a ‘view’ link under the Errors header but if a result is generated, it shows a ‘view’ link under the Results header. A ‘Refresh’ button can be clicked if the status of the query is not displayed on the screen.

Figure 102: Pathogenicity prediction in PolyPhen process 3

The view link displays the report of a queried variant predicting one of three status (probably damaging, possibly damaging or benign) with an associated score.

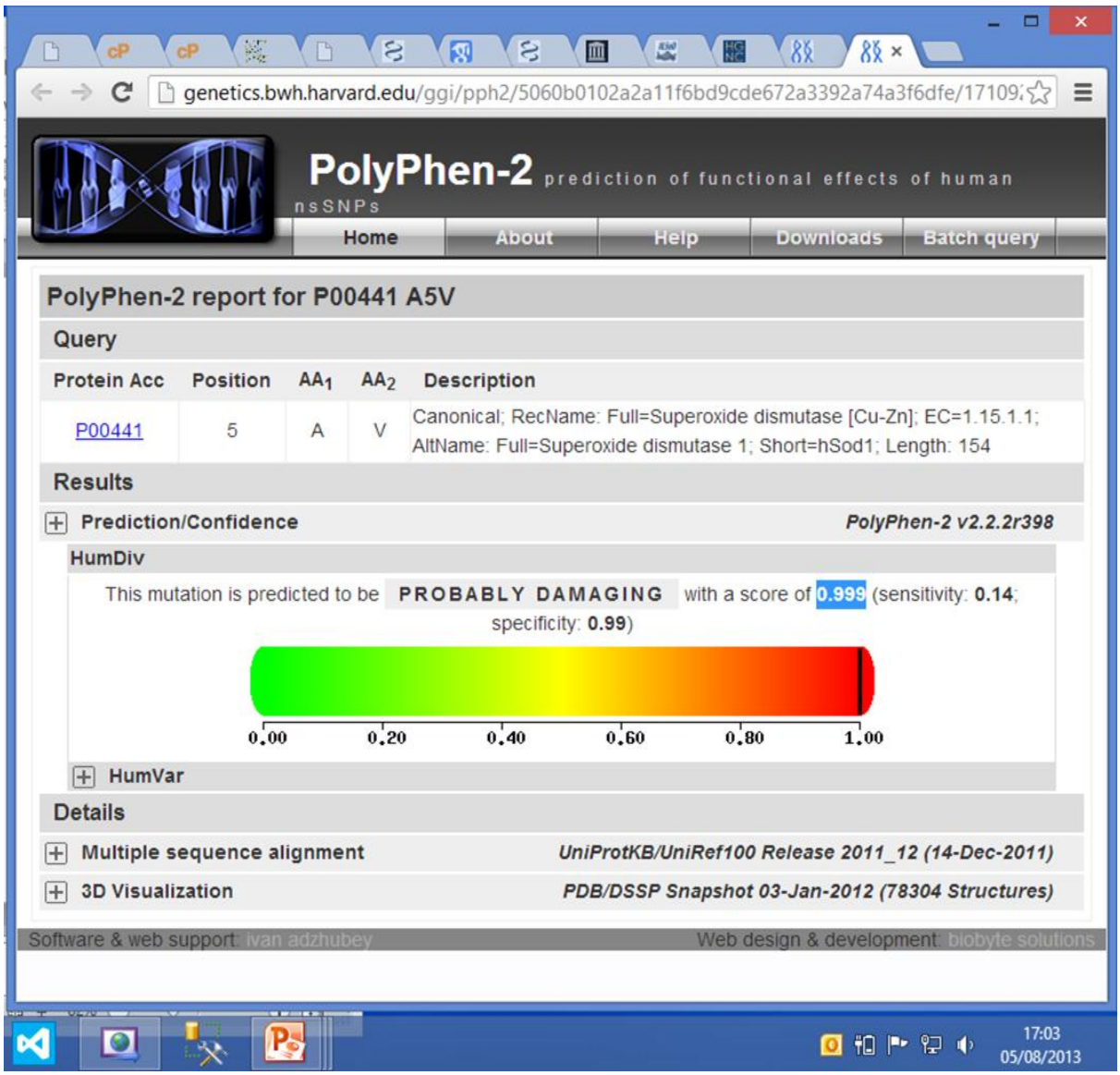


Figure 103: Pathogenicity prediction in PolyPhen process 4

6.4.2.5 Store data from generated analyses into database tables

A list of prediction text files with scores and predictions using the three bioinformatics tools (PANTHER, SIFT and POLYPHEN) on variants in genes are stored in a folder.

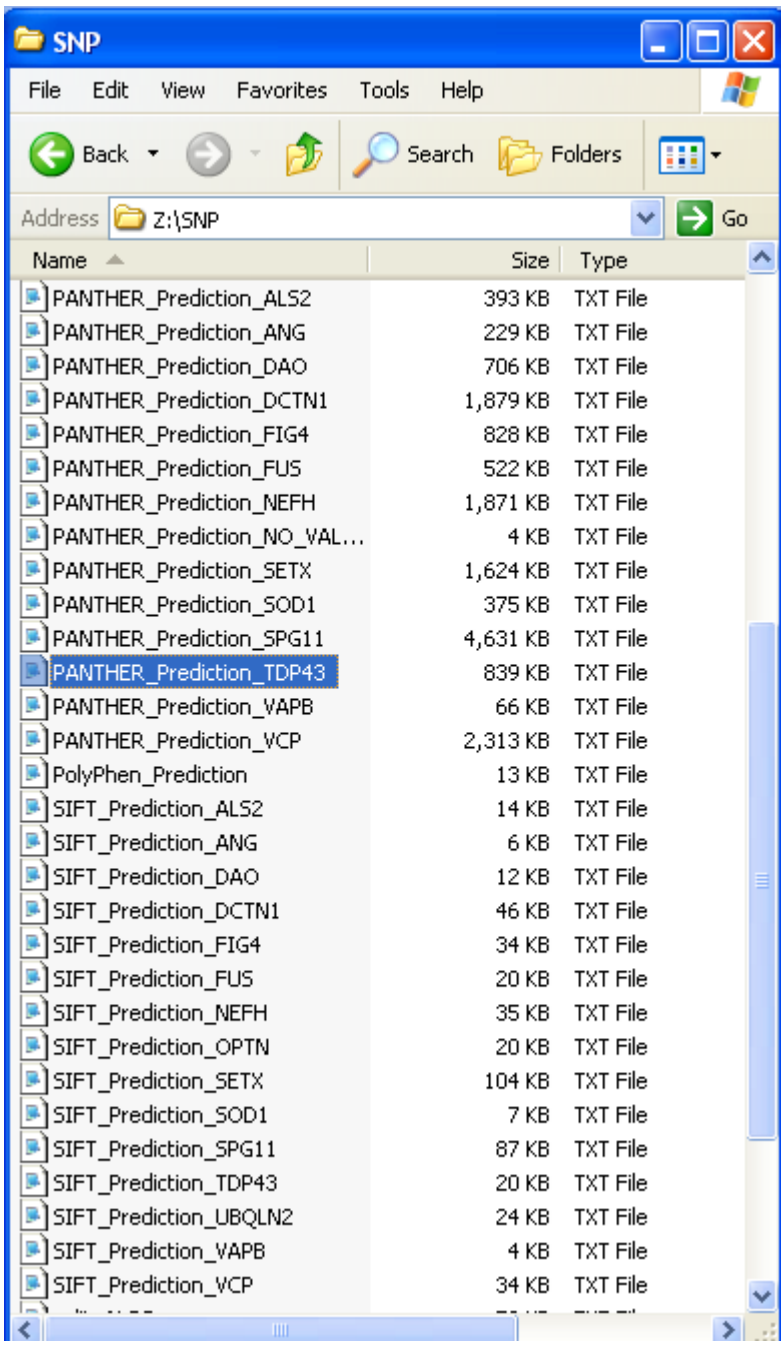


Figure 104: Storing data from generated analyses into database tables process 1

T-SQL scripts are written to create tables and a 'Bulk Insert' function is used to import the text files into the tables created. Column data automatically generated from the prediction tools are separated by 'tabs' which I used my discretion to select fields that are included in the tables.

dbo.[PANTHER_scores] table has 13 columns (subPSEC, P_deletrious, gene, mutation, codon, name, Accession, score, position, Pwt, Psubstituted, NIC, message)

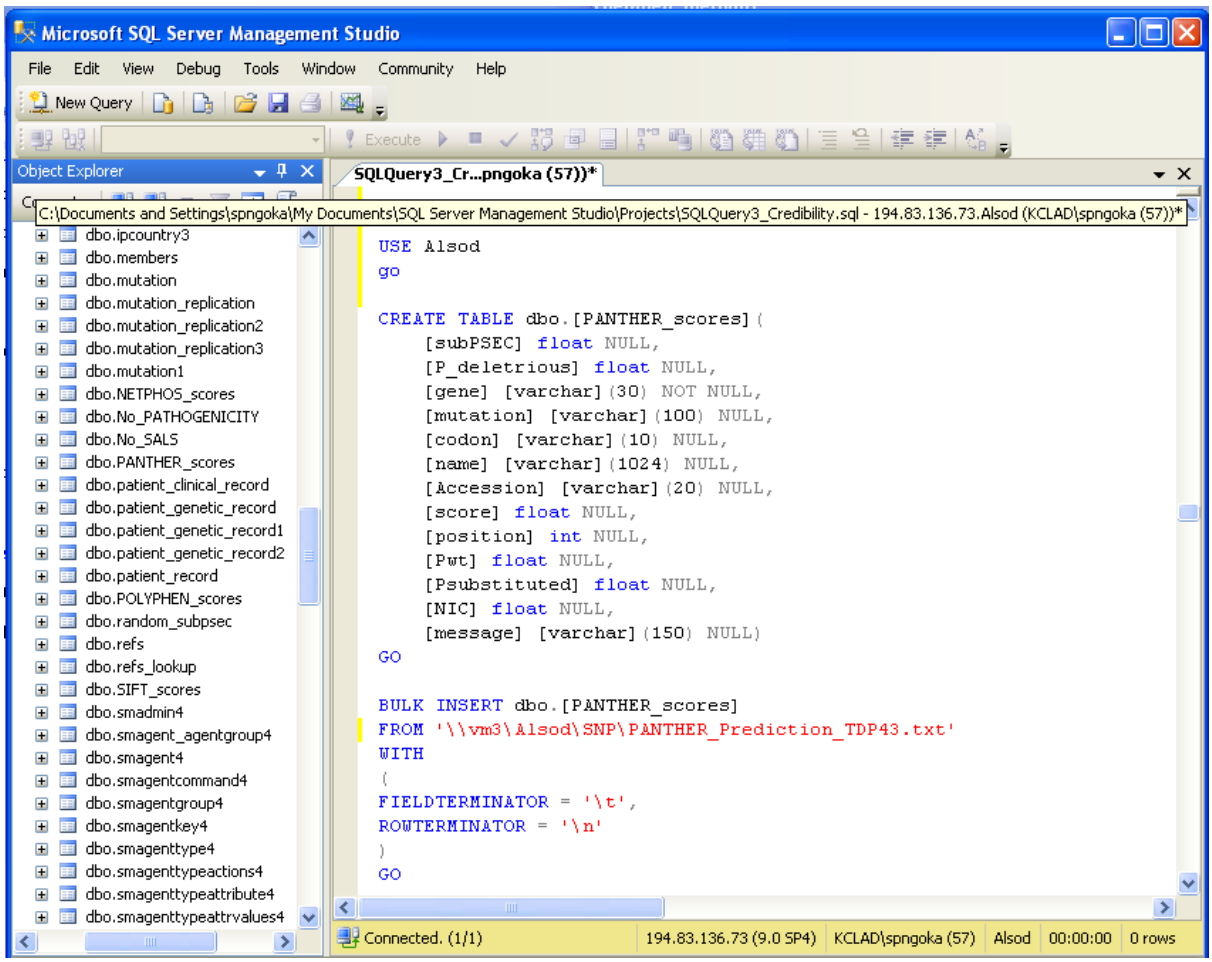


Figure 105: Storing data from generated analyses into database tables process 2

dbo.[SIFT_scores] table has 7 columns (position, codon, gene, mutation_from, mutation_to, message, subPSEC)

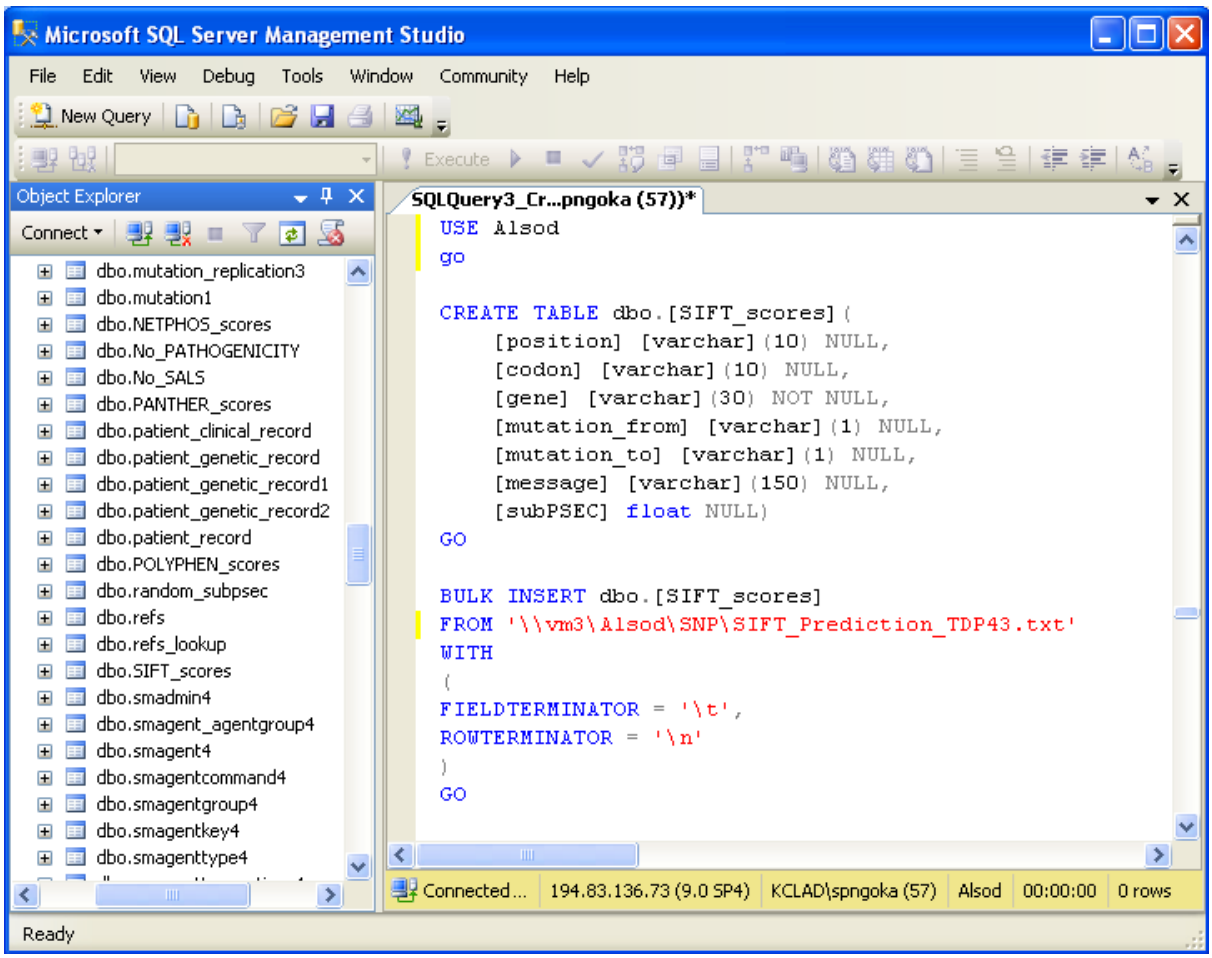


Figure 106: Storing data from generated analyses into database tables process 3

dbo.[PolyPhen_scores] table has 6 columns (gene, mutation, codon, position, PSIC_score, prediction)

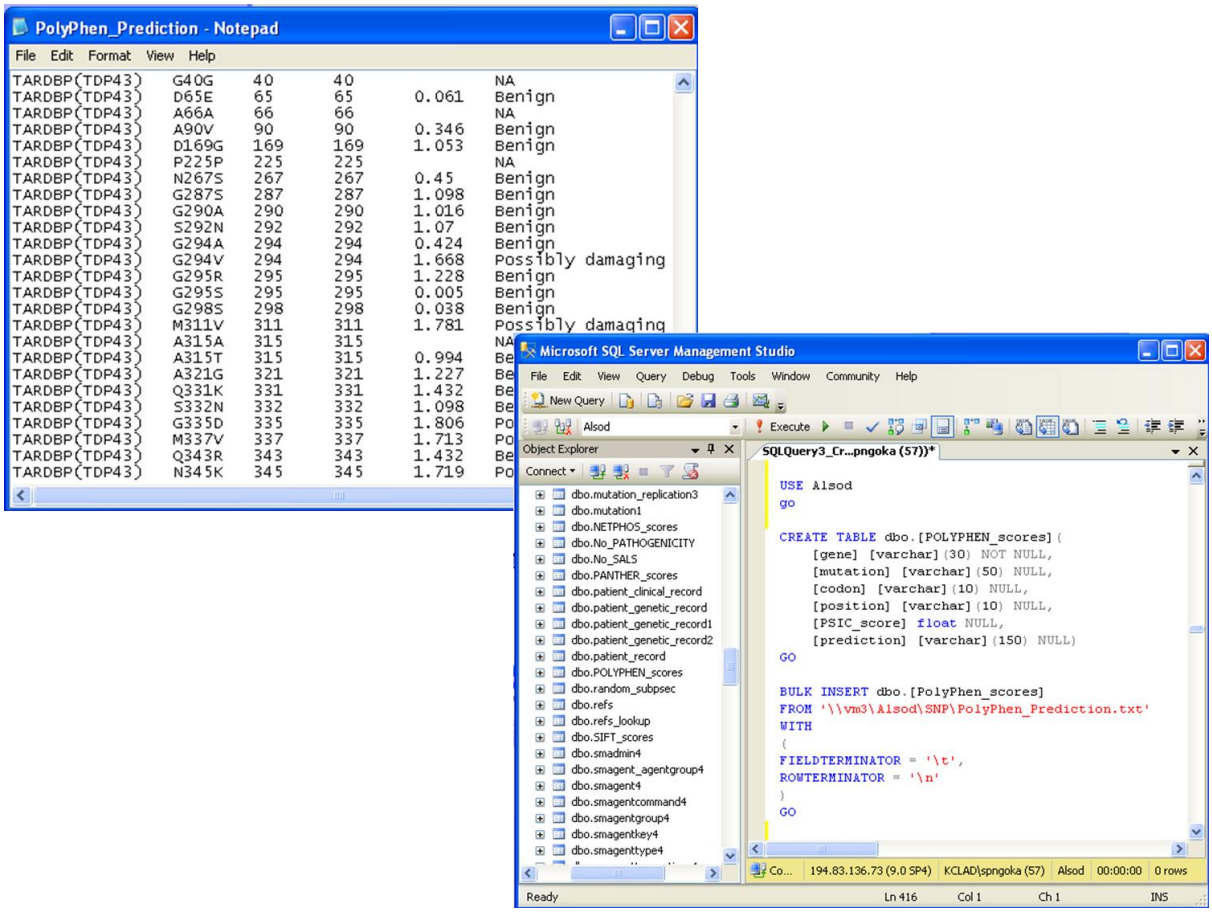


Figure 107: Storing data from generated analyses into database tables process 4

6.4.2.6 Query database on 3 tables for scores

The flowchart for querying the database is shown in Figure 108.

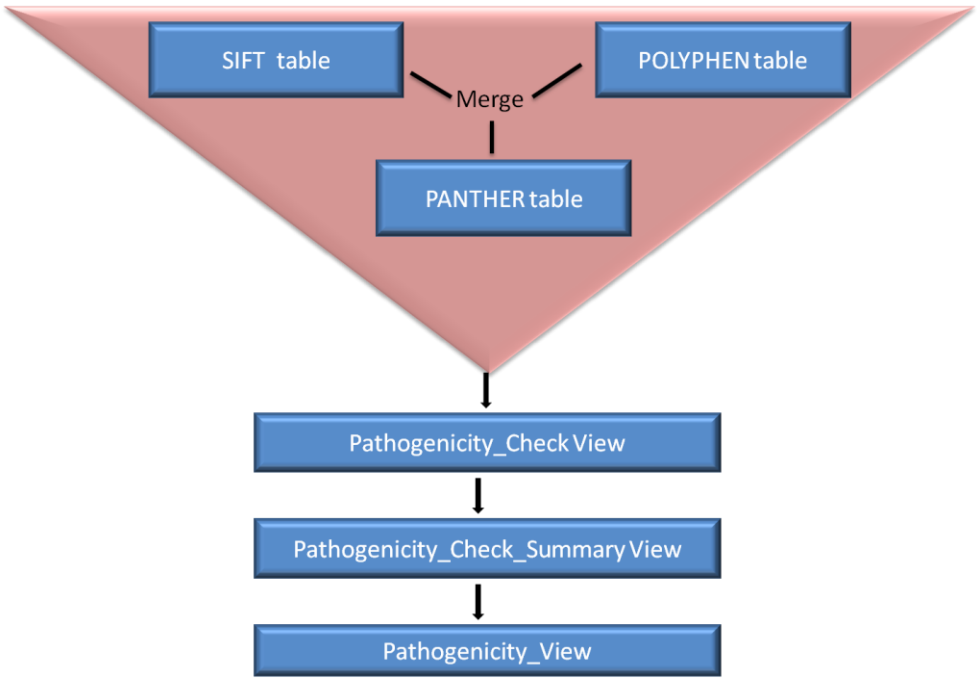


Figure 108: Flowchart of pathogenicity prediction query

```
SELECT     TOP (100) PERCENT dbo.PANTHER_scores.gene,
dbo.PANTHER_scores.mutation, dbo.PANTHER_scores.codon,
                      dbo.PANTHER_scores.subPSEC AS PANTHER,
dbo.PANTHER_scores.message AS Prediction1,
                      dbo.POLYPHEN_scores.PSIC_score_new AS POLYPHEN,
dbo.POLYPHEN_scores.prediction AS Prediction2, dbo.SIFT_scores.subPSEC AS SIFT,
                      dbo.SIFT_scores.message AS Prediction3, CASE WHEN
dbo.PANTHER_scores.message = 'Deletrious' THEN '1' ELSE '0' END AS
Panther_rate,
                      CASE WHEN (dbo.POLYPHEN_scores.prediction = 'Benign' OR
                      dbo.POLYPHEN_scores.prediction = 'NA') THEN '0' ELSE '1'
END AS Polyphen_rate,
                      CASE WHEN dbo.SIFT_scores.message = 'AFFECT PROTEIN
FUNCTION ' THEN '1' ELSE '0' END AS Sift_rate
FROM         dbo.PANTHER_scores INNER JOIN
                      dbo.POLYPHEN_scores ON dbo.PANTHER_scores.mutation =
dbo.POLYPHEN_scores.mutation AND
                      dbo.PANTHER_scores.gene = dbo.POLYPHEN_scores.gene INNER
JOIN
                      dbo.SIFT_scores ON dbo.POLYPHEN_scores.codon =
dbo.SIFT_scores.codon AND dbo.POLYPHEN_scores.gene = dbo.SIFT_scores.gene
```

```
SELECT     gene, mutation, codon, PANTHER, Prediction1, POLYPHEN, Prediction2,
SIFT, Prediction3, Panther_rate, Polyphen_rate, Sift_rate,
CAST(Panther_rate AS int) + CAST(Sift_rate AS int) + CAST(Polyphen_rate AS int)
AS Result, CASE WHEN (CAST(Panther_rate AS int) + CAST(Sift_rate AS int) +
CAST(Polyphen_rate AS int)) >= '1' THEN 'Yes' ELSE 'No' END AS Pathogenic
FROM         [KCLAD\spngoka].Pathogenicity_Check
```

```
SELECT     gene, COUNT(gene) AS Pathogenic
FROM         (SELECT     gene, mutation, Panther_rate, Sift_rate,
Polyphen_rate, Result, Pathogenic FROM
[KCLAD\spngoka].Pathogenicity_Check_Summary
WHERE      (Pathogenic = 'Yes')) AS derivedtbl_1 GROUP BY gene
```

6.4.3 Data extraction from publications

I conducted a systematic review of all publications related to ALS genetics with an exhaustive combination of search queries on the 15 genes mentioned above. (Review_protocol in Appendix 1, Flow_diagram in Appendix 6).

In the PubMed database, I used title keywords consisting of the gene name, "mutation" and "ALS" or “Motor Neuron Disease”, or gene name and "novel" to identify key publications and then used the related

citations function to generate a list of publications for data extraction. For example, (SOD1[Title] OR (superoxide dismutase[Title]) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) yielding 181 results. These results were further filtered by choosing “Humans” as Species and sorted by “Recently Added” thereby displaying 160 unique publications. From the list displayed, I also searched the “Related citations” link on the first publication[91] of the selected gene SOD1 yielding 204 results.

I used Google Scholar (http://scholar.google.co.uk/) to identify publications for import into the ALSoD database, starting with basic search queries to generate a large number of publications. For example, “SOD1” gave about 28,600 results but “SOD1 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease"” gave 2050 results. I went through the first 20 pages containing 20 publications on each page and already sorted by relevance. Publications with animal models or associated with other diseases were excluded from the long list. A manual comparison with already discovered publications from pubmed was conducted and these were excluded from the list.

Manually curated data extracted from all publications included family history, El Escorial category [35, 36] mutations per gene, number of cases and controls used in the studies, mutations in the same codon, number of patients with family history (FALS), number of patients without family history, mutations replicated in other studies, number of countries replicating the mutation and for linkage studies, LOD scores. Several genes implicated in ALS are also implicated in other diseases, including frontotemporal dementia, spinocerebellar ataxia and parkinsonism. To avoid the problem of non-ALS patients being included in the database, I restricted data curation to publications specifying ALS.

6.4.4 Automated gene ranking

Eleven queries stored as procedures were performed on data collated. These were: 1. The total number of affected patients with El Escorial defined ALS having a mutation in each gene [570]. 2. The total number of ALS affected patients used in each study. This measure was used to account for sampling variance and power [571]. 3. The total number of healthy individuals with a mutation reported in each study. 4. The total number of healthy individuals used in each study. 5. The total number of mutations sharing the same codon. 6. The total number of variants detected in ALS patients for each gene. 7. The total number of mutations with positive pathogenic predictions from the use of the three bioinformatics tools described above. 8. The number of patients with a family history defined as at least one other affected member of the family. 9. The number of patients without a family history of ALS. 10. The number of times a particular

variation was replicated across different studies. 11. The number of unique populations where affected patients originated.

6.4.4.1 Selecting criteria for analyses

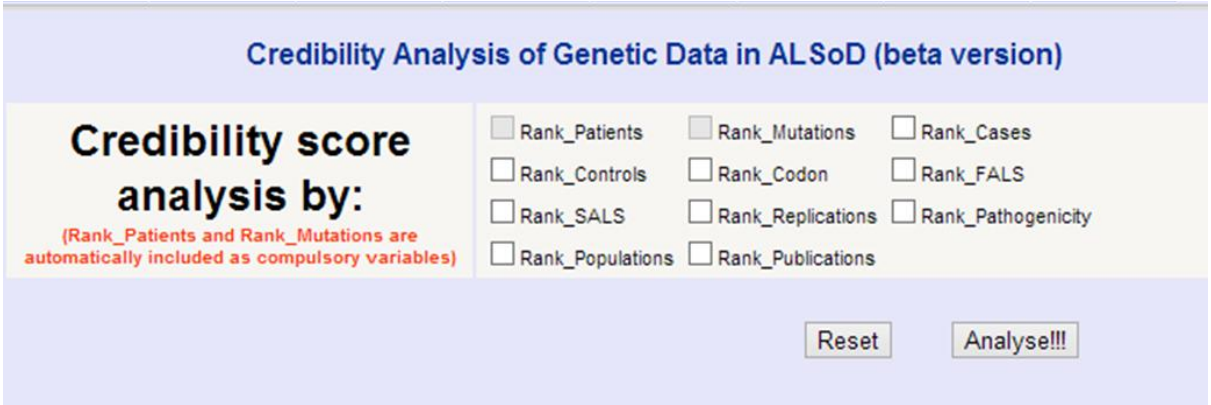


Figure 109: Criteria for Credibility score analysis

The criteria are checkboxes and users are asked to click the checkboxes next to the criteria they want to include in the analysis. The Rank_Patients and Rank_Mutations are automatically selected and used as compulsory variable to query the database. This is why they are greyed out.

```
<asp:CheckBoxList ID="tables" runat="server" AutoPostBack ="false"
 ToolTip="Select checkbox if desired as one of the criteria for analysing credibility score" RepeatColumns="3" RepeatDirection="Horizontal">
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_affected_patients_in_ALSoD]" Text="Rank_Patients" Enabled="false" >
</asp:ListItem>
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_mutations_per_gene]" Text="Rank_Mutations" Enabled="false">
</asp:ListItem>
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_cases_recorded]" Text="Rank_Cases">
</asp:ListItem>
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_controls_recorded]" Text="Rank_Controls">
</asp:ListItem>
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_mutations_in_same_codon_by_rank]" Text="Rank_Codon">
</asp:ListItem>
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_patients_with_family_history_FALS]" Text="Rank_FALS">
</asp:ListItem>
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_patients_without_family_history_SALS]" Text="Rank_SALS">
</asp:ListItem>
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_times_mutation_is_replicated]" Text="Rank_Replications">
</asp:ListItem>
<asp:ListItem Value="[KCLAD\spngoka].[Number_of_pathogenic_mutations_by_rank]" Text="Rank_Pathogenicity">
```

</asp:ListItem>

<asp:ListItem Value="[KCLAD\spngoka].[Number_of_unique_countries_on_genes]" Text="Rank_Populations">

</asp:ListItem>

</asp:CheckBoxList>

6.4.4.2 Data display on webpage

<h2> Details of ranked credibility data </h2>

1) Number_of_affected_patients_in_ALSoD

2) Number_of_mutations_per_gene

3) Number_of_cases_recorded

4) Number_of_predicted_pathogenic_mutations_by_rank

5) Number_of_controls_recorded

6) Number_of_mutations_in _same_codon_by_rank

7) Number_of_patients_with_family_history_FALS

8) Number_of_patients_with_no_family_history_SALS

9) Number_of_times_mutation_is_replicated

10) Number_of_unique_countries_on_genes

6.4.4.3 Query generation in SQL

For each procedure above, a query was generated using Structured Query Language (SQL) on Microsoft SQL Server 2008 and displayed on the ASP.NET platform webpage, ranking the gene.

6.4.4.3.1 Number_of_affected_patients_in_ALSoD

SELECT [Gene], [Frequency] As Patients, [Rank] FROM [KCLAD\spngoka].[Number_of_affected_patients_in_ALSoD]

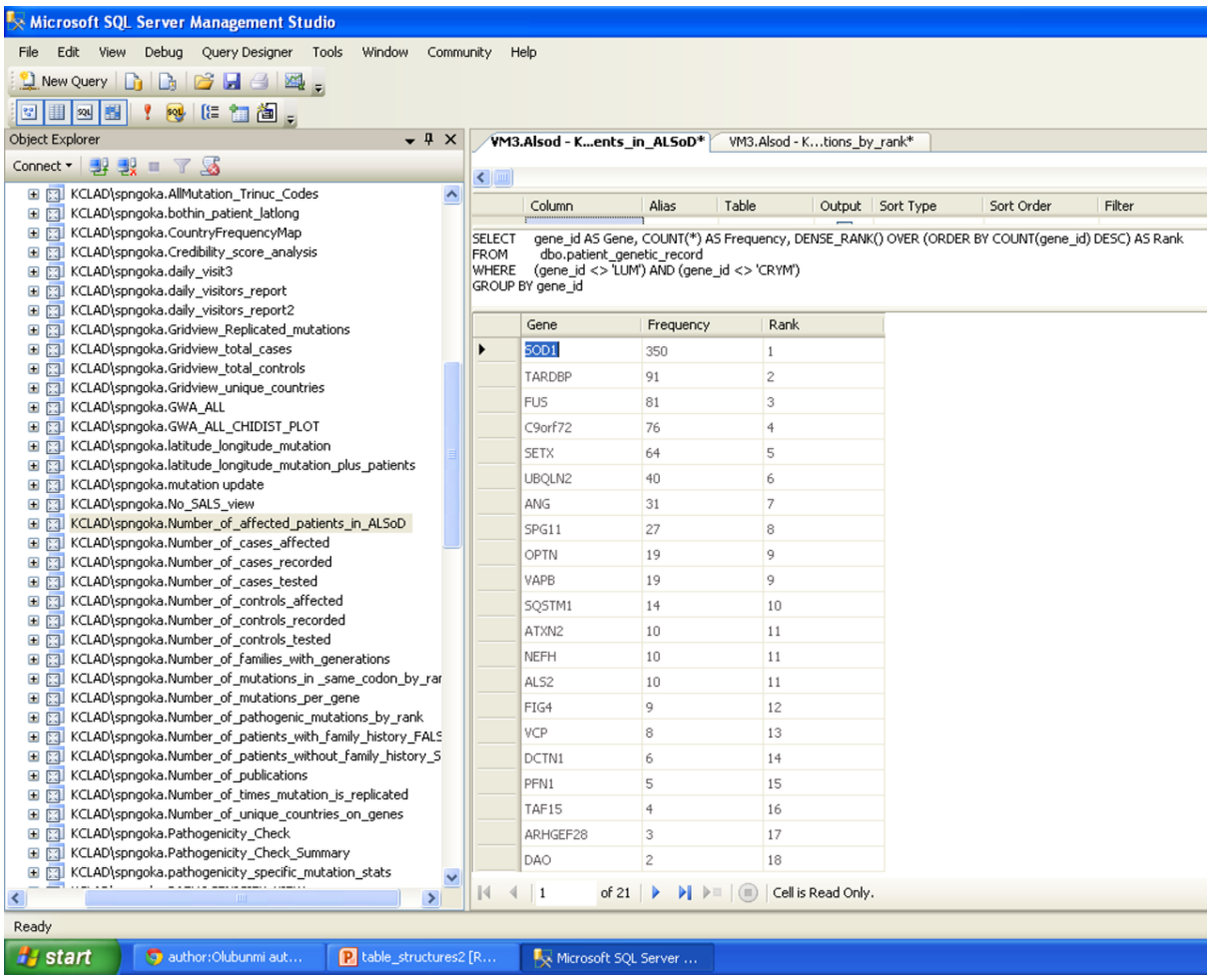


Figure 110: Number of affected patients in ALSoD

6.4.4.3.2 Number_of_mutations_per_gene

SELECT [Gene], [Frequency] As Mutations, [Rank] FROM [KCLAD\spngoka].[Number_of_mutations_per_gene]

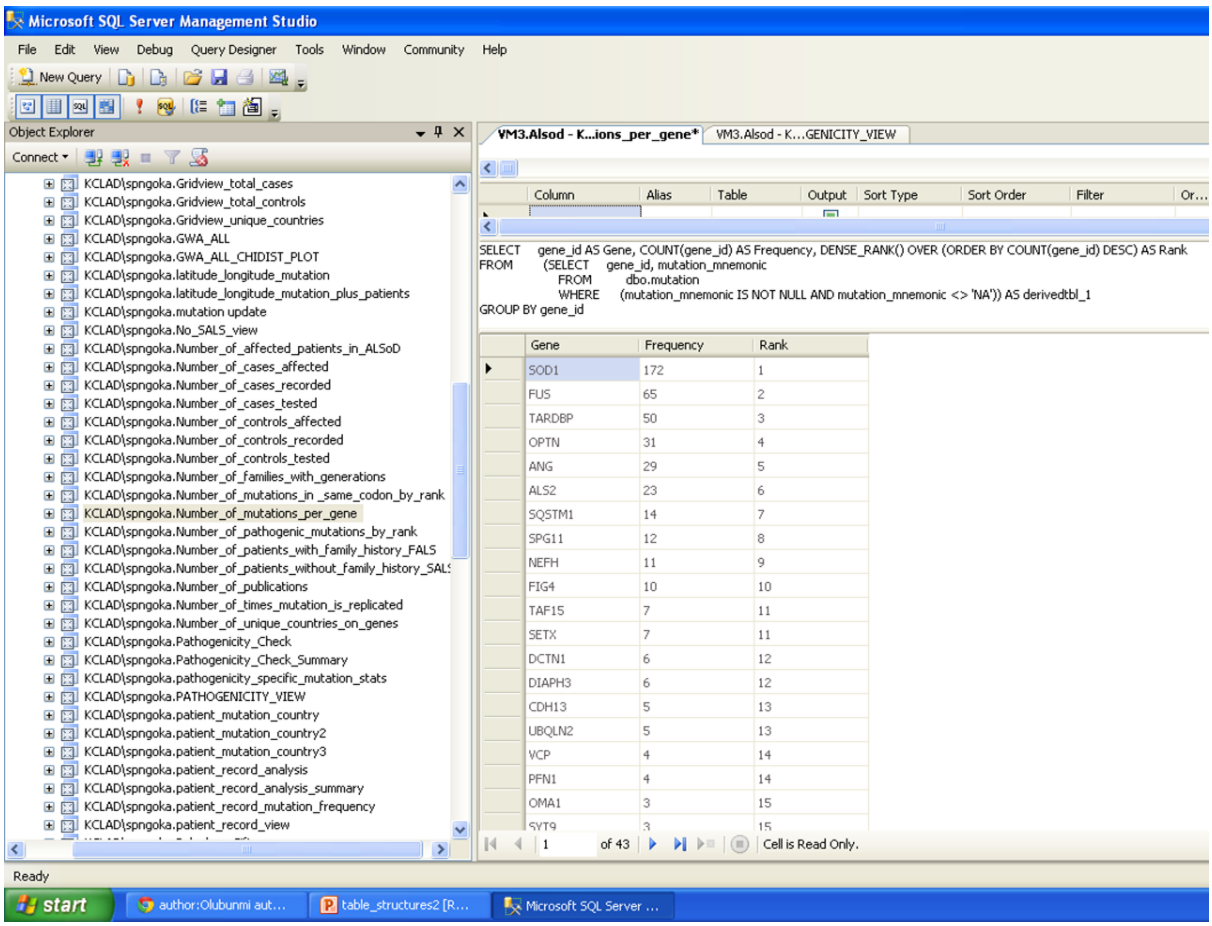


Figure 111: Number of mutations per gene

6.4.4.3.3 Number_of_cases_recorded

SELECT DISTINCT [gene], [cases], [author], [year], [pubmed_id] FROM [KCLAD\spngoka].[Gridview_total_cases]

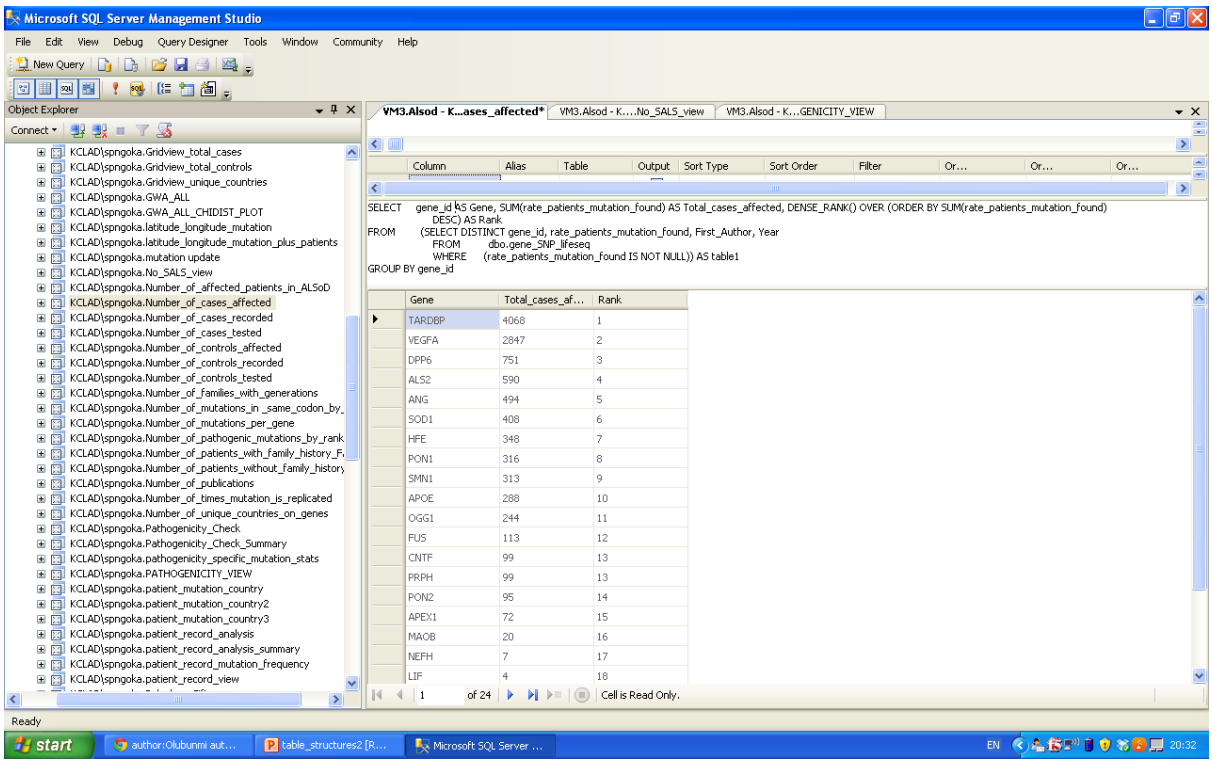


Figure 112: Number of cases recorded

6.4.4.3.4 Number_of_predicted_pathogenic_mutations_by_rank

SELECT [Gene], [Total_cases_affected], [Rank] FROM [KCLAD\spngoka].[Number_of_cases_affected]

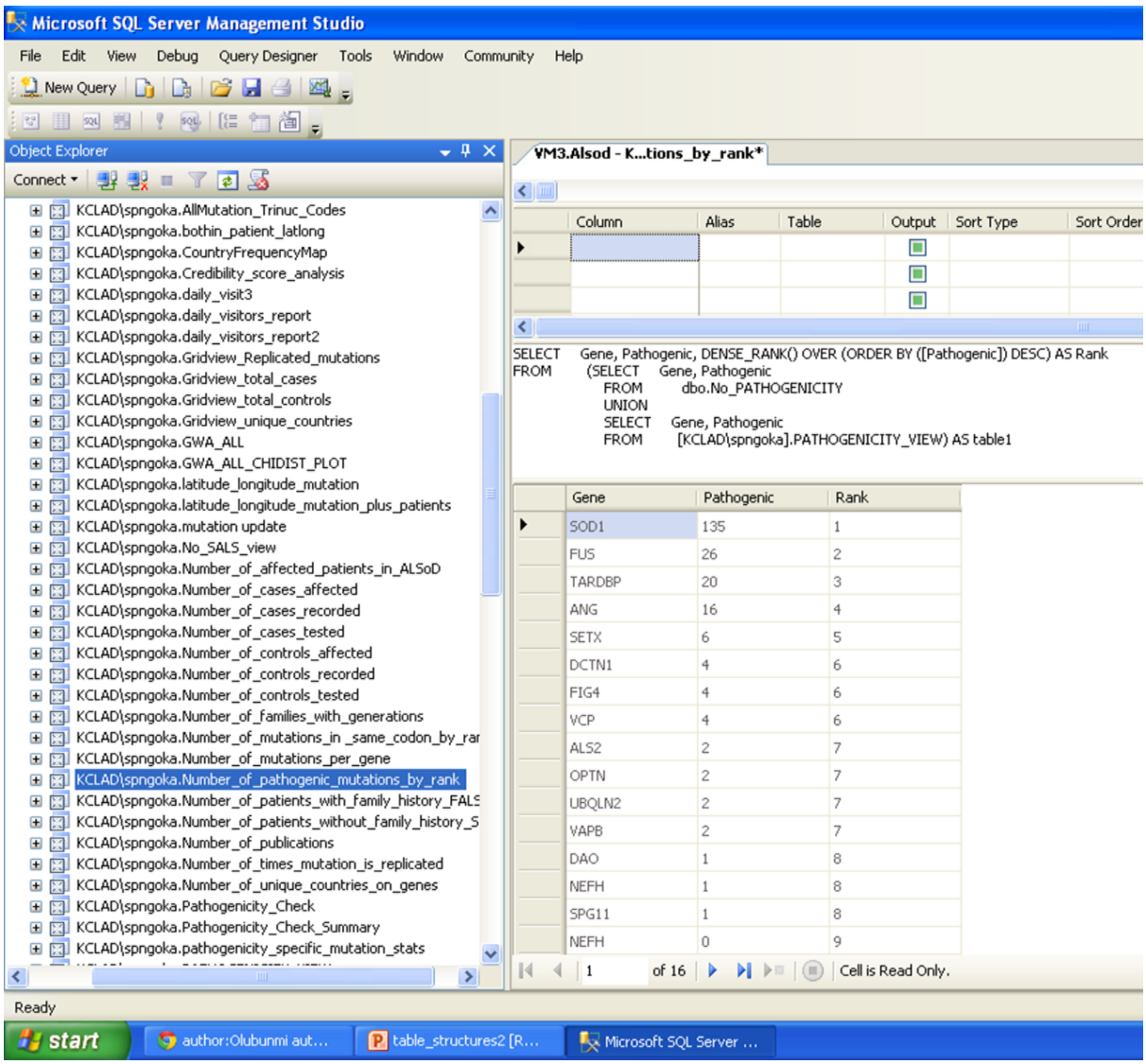


Figure 113: Number of predicted pathogenic mutations by rank

6.4.4.3.5 Number_of_controls_recorded

SELECT [Gene], [Total_controls_affected], [Rank] FROM [KCLAD\spngoka].[Number_of_controls_affected]

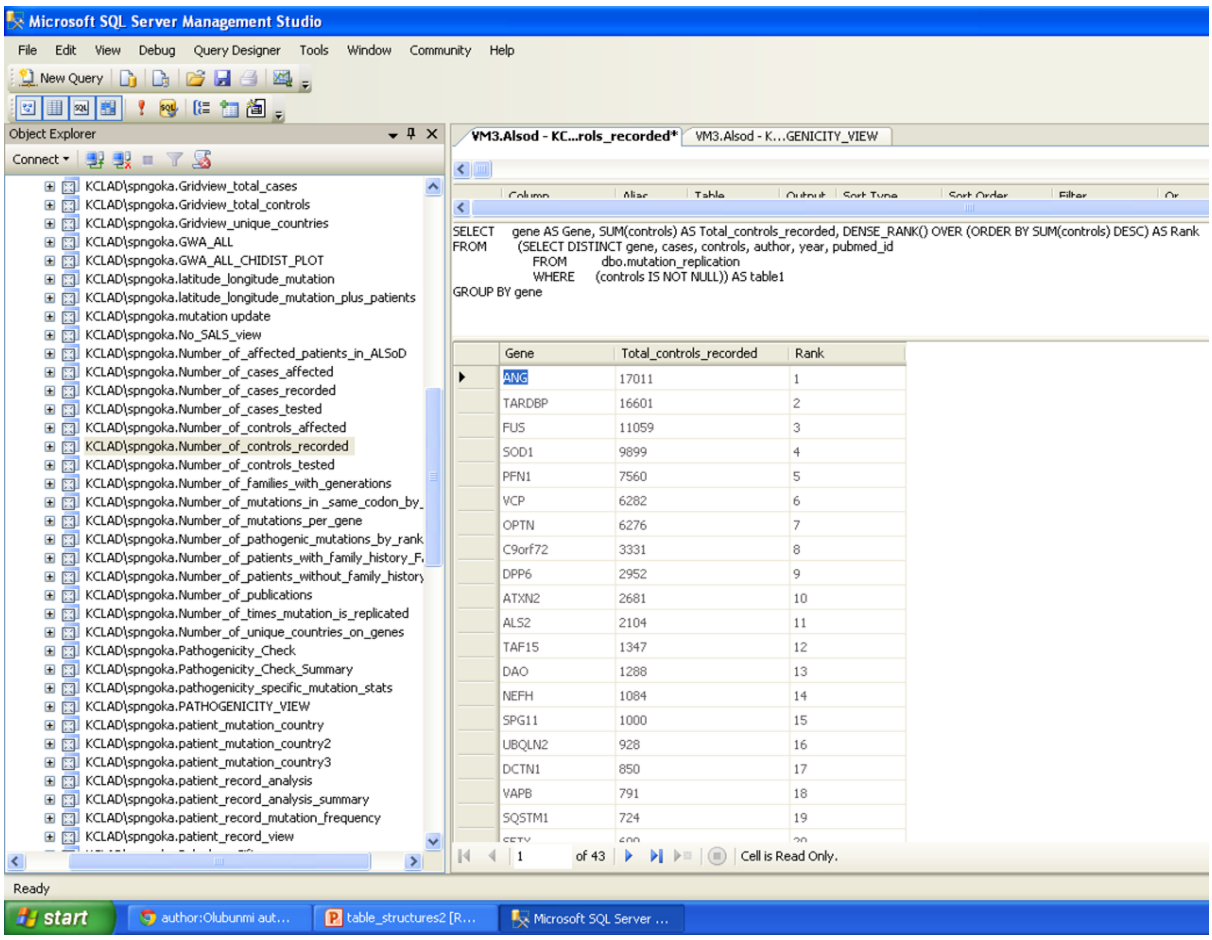


Figure 114: Number of controls recorded

6.4.4.3.6 Number_of_mutations_in _same_codon_by_rank

SELECT gene_id AS Gene, codon AS Codon, COUNT(codon) AS Frequency, DENSE_RANK() OVER (ORDER BY COUNT(codon) DESC) AS Rank FROM (SELECT TOP (100) PERCENT gene_id, mutation_mnemonic, codon FROM dbo.mutation WHERE (codon IS NOT NULL) AND (codon <> 0) ORDER BY codon) AS table1 GROUP BY gene_id, codon ORDER BY Gene

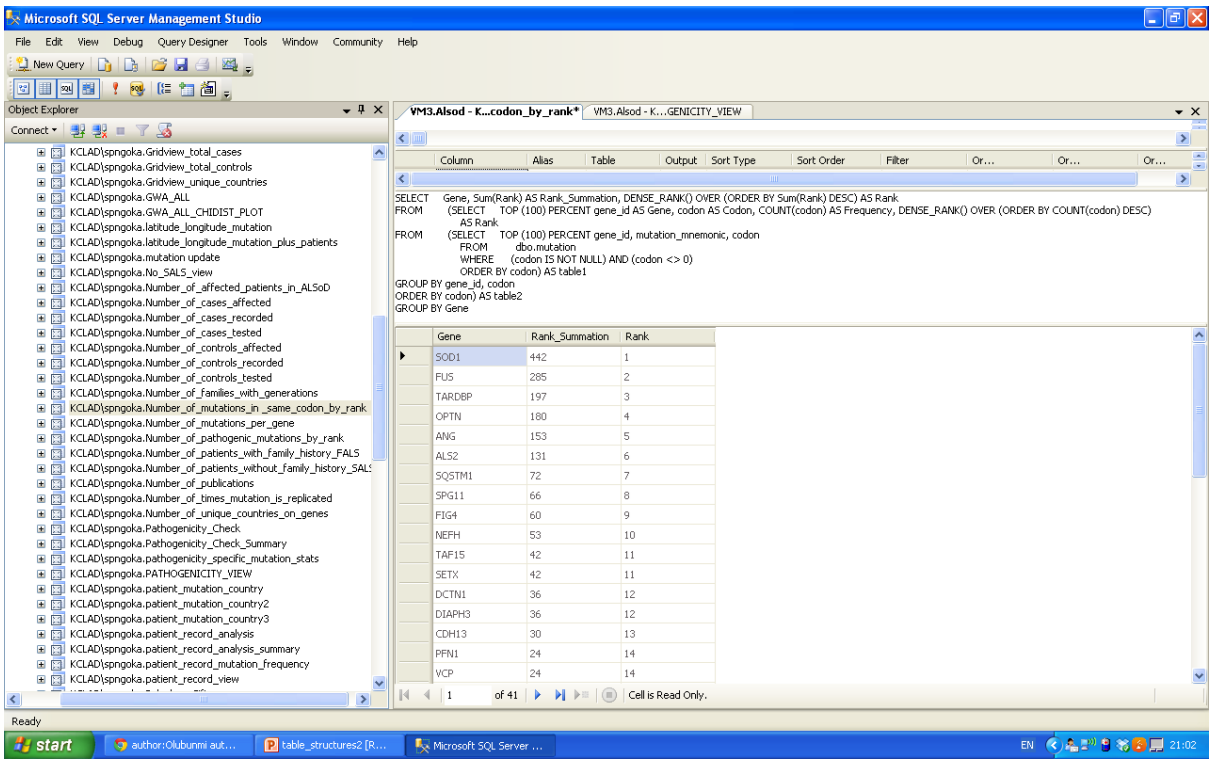


Figure 115: Number of mutations in same codon by rank

6.4.4.3.7 Number_of_patients_with_family_history_FALS

SELECT [Gene], [Total_family_history], [Rank] FROM [KCLAD\spngoka].[Number_of_patients_with_family_history_FALS]

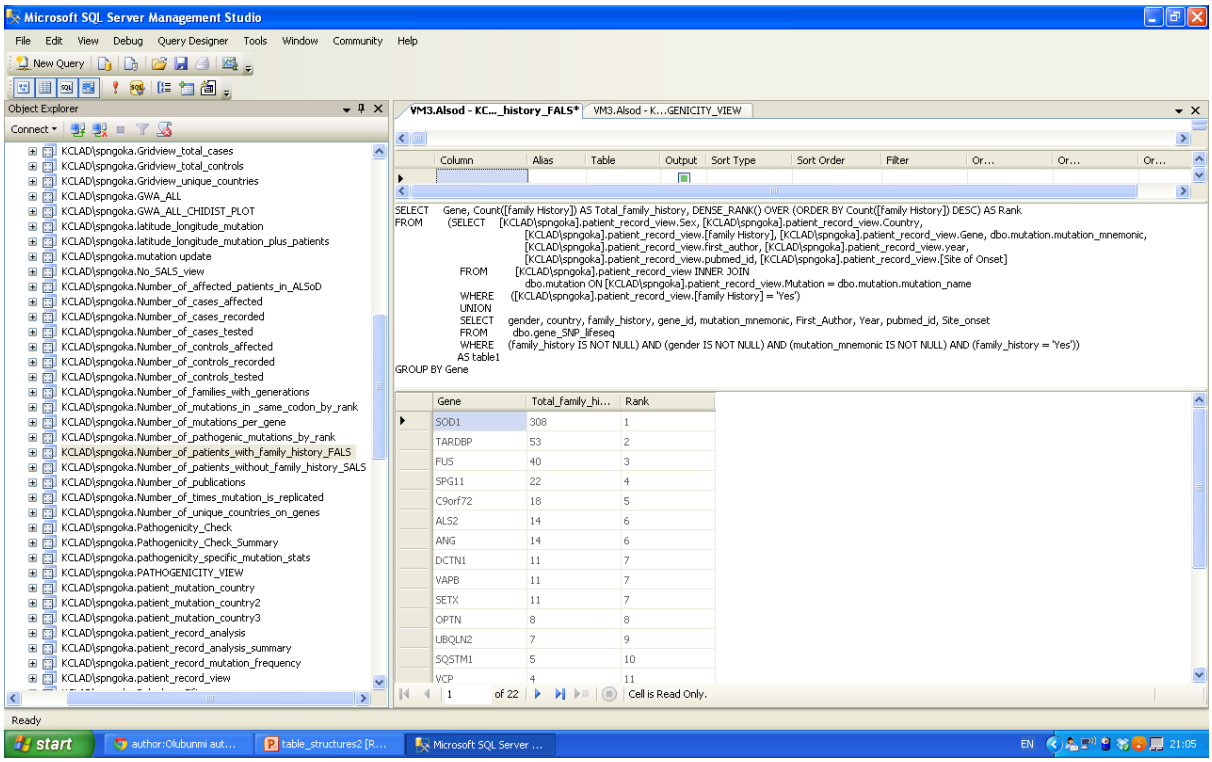


Figure 116: Number of patients with family history FALS

6.4.4.3.8 Number_of_patients_with_no_family_history_SALS

SELECT [Gene], [Total_no_family_history], [Rank] FROM [KCLAD\spngoka].[Number_of_patients_without_family_history_SALS]

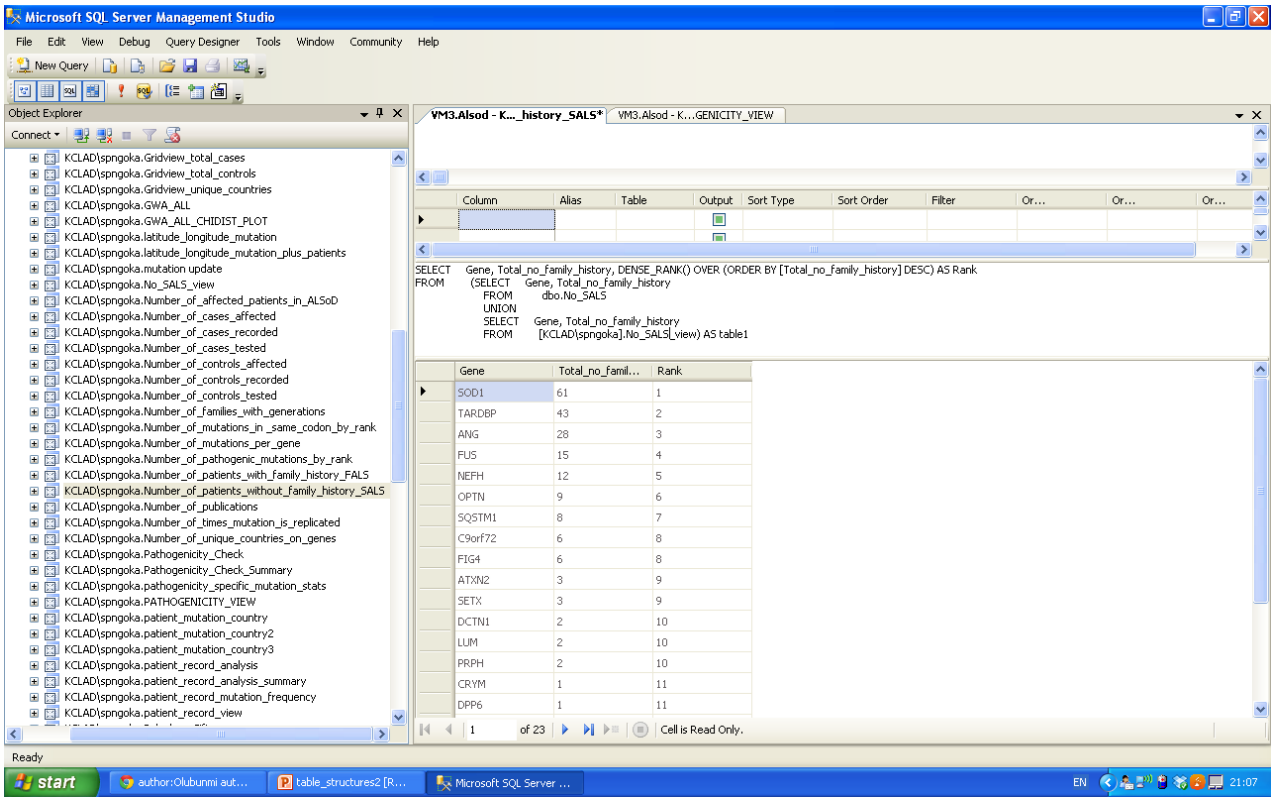


Figure 117: Number of ptients with SALS

6.4.4.3.9 Number_of_times_mutation_is_replicated

SELECT [Gene], [Mutation], [Frequency], [Rank_Mutation] FROM [KCLAD\spngoka].[Gridview_Replicated_mutations] WHERE ([Frequency] >= @Frequency)

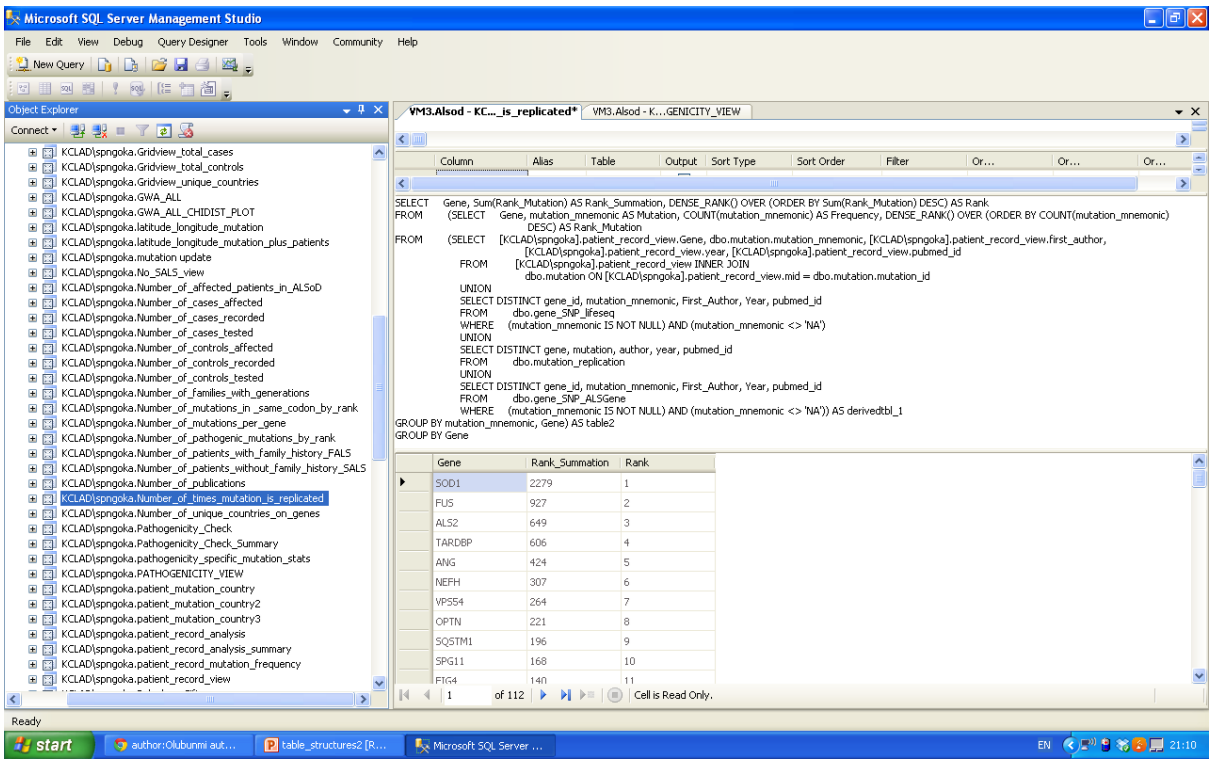


Figure 118: Number of times mutation is replicated

6.4.4.3.10 Number_of_unique_countries_on_genes

SELECT [gene_id], [country] FROM [KCLAD\spngoka].[Gridview_unique_countries]

SELECT [Gene], [Total_Countries], [Rank] FROM [KCLAD\spngoka].[Number_of_unique_countries_on_genes]

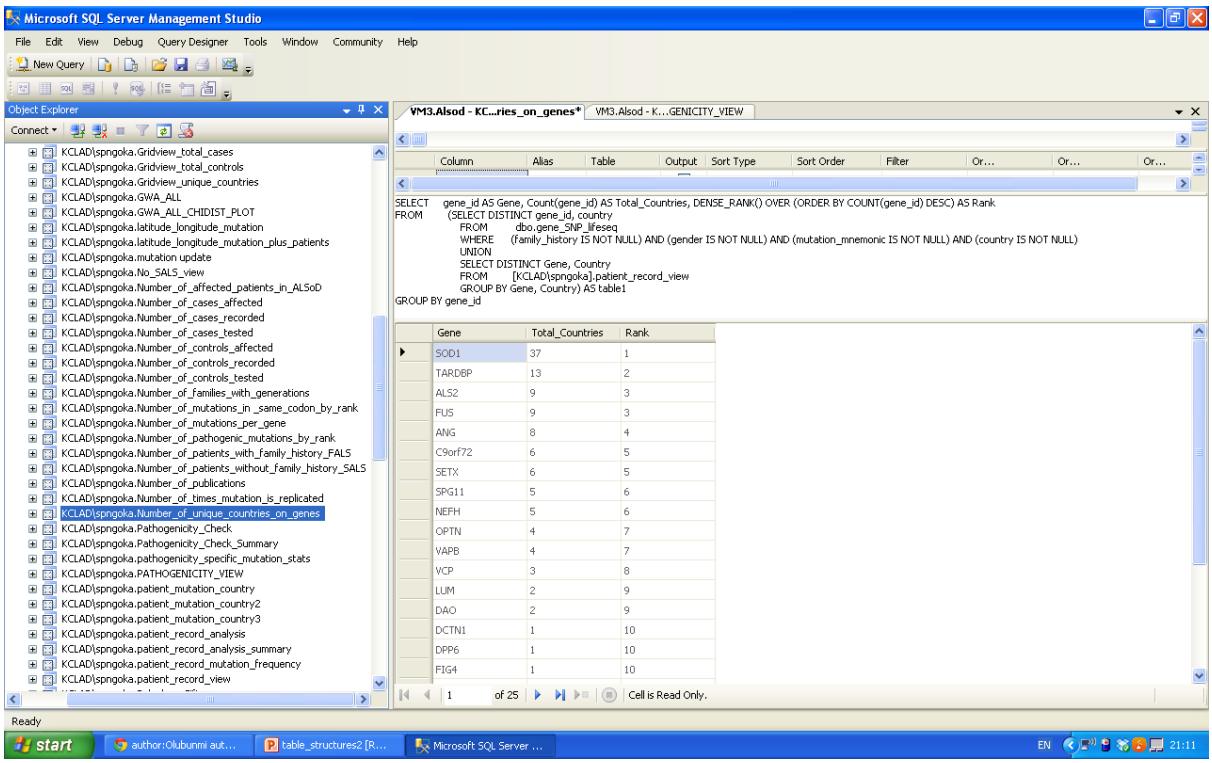


Figure 119: Number of unique countries on genes

6.4.4.4 Ranking criteria

The predicted pathogenicity score for each tool was scored 1 for predicted pathogenic and 0 for predicted not pathogenic and then summed to generate a final score for ranking (http://alsod.iop.kcl.ac.uk/Statistics/pathogenicity.aspx).

There are two possible ways of ranking results in SQL.

6.4.4.4.1 RANK( )

The default method allocates rank based on the true position, such that if two genes are given equal first position for example, the next gene is in third position, not second.

6.4.4.4.2 DENSE RANK( )

The dense rank method allocates the next gene as second so that there are no gaps in the rank numbering. I used the dense ranking system.

6.4.4.5 Testing

I started planning on how to derive a formular to run an SQL script depending on the checked criteria. So, on SQL, I included all the criteria to get a lengthy script as shown in Appendix 21.

I later ran an SQL on two criteria (rank_mutation and rank_patient) to compare with the first step (selecting all criteria). The script is shown below:

```
SELECT Rank_Mutations, Rank_Patients, Gene ,SUM( Rank_Mutations + Rank_Patients + 0) As Rank_Sum, DENSE_RANK() OVER (ORDER BY SUM( Rank_Mutations + Rank_Patients + 0) ASC) AS Final_Rank FROM (SELECT [KCLAD\spngoka].Number_of_mutations_per_gene.Rank As Rank_Mutations,[KCLAD\spngoka].Number_of_affected_patients_in_ALSoD.Rank AS Rank_Patients, [KCLAD\spngoka].[Number_of_mutations_per_gene].Gene FROM [KCLAD\spngoka].Number_of_affected_patients_in_ALSoD INNER JOIN [KCLAD\spngoka].Number_of_mutations_per_gene ON [KCLAD\spngoka].Number_of_affected_patients_in_ALSoD.Gene = [KCLAD\spngoka].Number_of_mutations_per_gene.Gene ) As table1 GROUP BY Rank_Mutations, Rank_Patients, Gene
```

A formula was derived as shown below:

SELECT All_Rank_Criteria, Gene, SUM (Total (All_Rank_Criteria)) As Rank_sum,

DENSE_RANK() Over (ORDER BY SUM (Total(All_Rank_Criteria)ASC) AS Final_Rank

FROM (SELECT [Table*.*.Rank As All_Rank_Criteria] FROM CONCATENATE

(Table*.* INNER JOIN Table.Number_of_mutations_per_gene ON Table*.*.Gene=

Table.Number_of_mutations_per_gene.Gene) As table1 GROUP BY All_Rank_Criteria, Gene

Where :

All_Rank_Criteria= Rank_Mutations, Rank_Patients, Rank_Cases, Rank_Controls, Rank_Codon, Rank_FALS, Rank_SALS, Rank_Replications, Rank_Pathogenicity, Rank_Populations, Gene

SUM (Total (All_Rank_Criteria) = Rank_Mutations + Rank_Patients + Rank_Cases + Rank_Controls + Rank_Codon + Rank_FALS + Rank_SALS + Rank_Replications + Rank_Pathogenicity + Rank_Populations

Table*.* = [KCLAD\spngoka].[Number_of_mutations_per_gene],

[KCLAD\spngoka].[Number_of_affected_patients_in_ALSoD],

[KCLAD\spngoka].[Number_of_cases_recorded],

[KCLAD\spngoka].[Number_of_controls_recorded],

[KCLAD\spngoka].[Number_of_mutations_in _same_codon_by_rank],

[KCLAD\spngoka].[Number_of_patients_with_family_history_FALS],

[KCLAD\spngoka].[Number_of_patients_without_family_history_SALS],

[KCLAD\spngoka].[Number_of_times_mutation_is_replicated],

[KCLAD\spngoka].[Number_of_pathogenic_mutations_by_rank],

[KCLAD\spngoka].[Number_of_unique_countries_on_genes],

[KCLAD\spngoka].[Number_of_mutations_per_gene]

The visibility of the labels on the website was set to false so that it would not be seen by users as shown in Appendix 25.

6.4.5 Validation of the method

The purpose of the credibility score tool is to generate a list of genes in order of the weight of evidence supporting involvement in ALS. Such a list should correlate closely with one generated by ALS genetics experts, since such experts should have a good working knowledge of the available evidence. We therefore conducted a survey of ALS genetic experts, defined as being individuals who had published as first or senior author on ALS genetics.

6.4.5.1 Survey webpage

I embedded the questionnaire as a submenu on the feedback menu of the ALSoD website. Survey filled in here are linked to survey monkey so that there could be a single result from the same port irrespective of where the survey is filled. The link is found under Feedback -> Survey with direct link to http://alsod.iop.kcl.ac.uk/database/gene/credibilitySurveymonkey.aspx

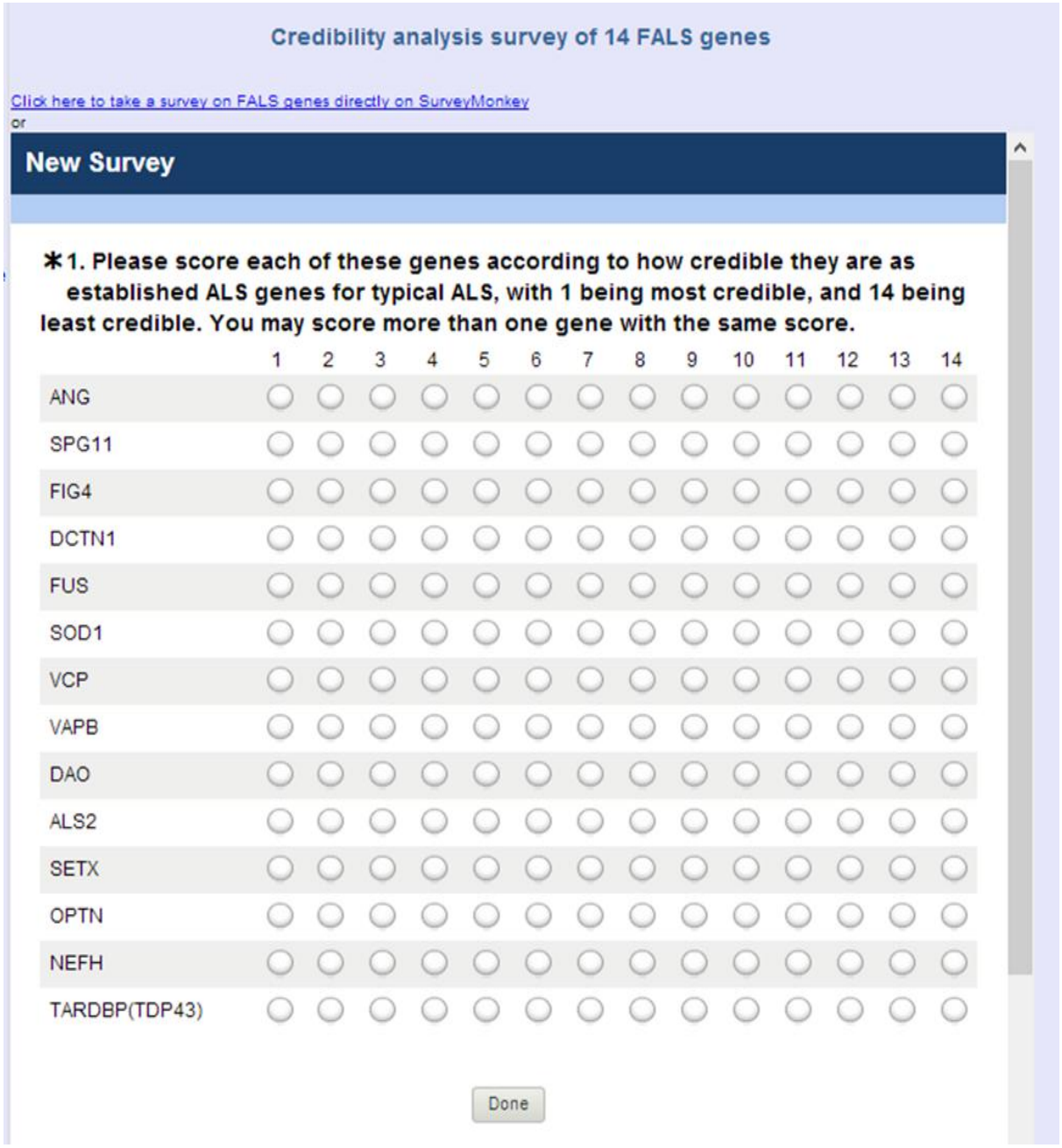


Figure 120: Survey displayed on webpage

6.4.5.2 Surveymonkey

Experts were surveyed using the freely available online questionnaire tool, Surveymonkey on http://www.surveymonkey.com/s/WRDW5WT (Figure 120). The survey link showed the genes randomly ordered differently every time the link was clicked to prevent bias in the responses that might occur based on ordering. Experts were randomly assigned to one of two groups, one in which the same rank could be assigned to several genes, and one in which responders were forced to rank each gene in order. The first group mimics the final score of the automated method closely, while the second group mimics the detail of the automated ranking method closely, since the automated method is forced to rank each query uniquely but the combined ranking could result in the same value for different genes.

6.4.5.2.1 Forced ranking

The forced ranking group have a list of the 14 genes and experts are asked to rank the 14 genes from 1st to 14th position. This type of ranking is selected while configuring the survey on the surveymonkey website.

ALS2	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	16.7% (1)	0.0% (0)	16.7% (1)	0.0% (0)	0.0% (0)	0.0% (0)	**33.3% (2)**	**33.3% (2)**	11.67	6
ANG	0.0% (0)	0.0% (0)	0.0% (0)	**33.3% (2)**	16.7% (1)	16.7% (1)	16.7% (1)	16.7% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	5.67	6
DAO	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	**33.3% (2)**	16.7% (1)	**33.3% (2)**	0.0% (0)	0.0% (0)	0.0% (0)	16.7% (1)	9.83	6
DCTN1	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	16.7% (1)	16.7% (1)	0.0% (0)	16.7% (1)	16.7% (1)	**33.3% (2)**	12.00	6
FIG4	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	16.7% (1)	0.0% (0)	0.0% (0)	0.0% (0)	**50.0% (3)**	16.7% (1)	16.7% (1)	0.0% (0)	10.83	6
FUS	16.7% (1)	33.3% (2)	**50.0% (3)**	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	2.33	6
NEFH	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	16.7% (1)	16.7% (1)	0.0% (0)	0.0% (0)	**33.3% (2)**	16.7% (1)	16.7% (1)	0.0% (0)	0.0% (0)	9.33	6
OPTN	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	**66.7% (4)**	33.3% (2)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	5.33	6
SETX	0.0% (0)	0.0% (0)	0.0% (0)	**16.7% (1)**	0.0% (0)	**16.7% (1)**	0.0% (0)	0.0% (0)	0.0% (0)	**16.7% (1)**	0.0% (0)	**16.7% (1)**	**16.7% (1)**	**16.7% (1)**	9.83	6
SPG11	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	16.7% (1)	0.0% (0)	16.7% (1)	0.0% (0)	**33.3% (2)**	16.7% (1)	16.7% (1)	0.0% (0)	10.50	6
SOD1	**50.0% (3)**	16.7% (1)	33.3% (2)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	1.83	6
TARDBP(TDP43)	33.3% (2)	**50.0% (3)**	16.7% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	1.83	6
VAPB	0.0% (0)	0.0% (0)	0.0% (0)	16.7% (1)	0.0% (0)	16.7% (1)	0.0% (0)	**33.3% (2)**	16.7% (1)	0.0% (0)	0.0% (0)	16.7% (1)	0.0% (0)	0.0% (0)	7.83	6
VCP	0.0% (0)	0.0% (0)	0.0% (0)	**33.3% (2)**	16.7% (1)	0.0% (0)	16.7% (1)	16.7% (1)	16.7% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	6.17	6

Figure 121: Forced ranking output

6.4.5.2.2 Non-forced ranking

The non-forced ranking group allows researchers to rank the 14 genes in the way that best suits them. So, more than one gene could be assigned the first position.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Rating Average	Response Count
ALS2	0.0% (0)	40.0% (2)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	40.0% (2)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	7.20	5
ANG	0.0% (0)	40.0% (2)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	6.20	5
DAO	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	40.0% (2)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	20.0% (1)	0.0% (0)	0.0% (0)	7.40	5
DCTN1	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	40.0% (2)	0.0% (0)	0.0% (0)	20.0% (1)	9.40	5
FIG4	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	40.0% (2)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	20.0% (1)	9.00	5
FUS	100.0% (5)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	1.00	5
NEFH	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	20.0% (1)	0.0% (0)	20.0% (1)	9.80	5
OPTN	0.0% (0)	20.0% (1)	60.0% (3)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	3.40	5
SETX	0.0% (0)	20.0% (1)	20.0% (1)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	6.60	5
SPG11	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	20.0% (1)	8.40	5
SOD1	100.0% (5)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	1.00	5
TARDBP(TDP43)	100.0% (5)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	1.00	5
VAPB	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	20.0% (1)	40.0% (2)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	6.20	5
VCP	40.0% (2)	40.0% (2)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	20.0% (1)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	0.0% (0)	2.60	5

Figure 122: Non-forced ranking

6.4.5.3 Emails

Emails were sent to 25 randomly selected experts in the field of ALS, inviting them to take part in a short survey. A sample of the letter is shown in Appendix 34.

6.4.6 User interface

The Credibility Analysis page at (http://alsod.iop.kcl.ac.uk/Statistics/credibility.aspx) has a list of 11 checkboxes which are the available choices of criteria a user can make. The code behind the scene for the checkboxes is as seen in Appendix 23.

6.4.7 Statistical methods

Cronbach's Alpha was initially used and then Spearman's Rho [572-574] was used to compare rankings generated by the automated method and the ALS genetics experts. We agreed that Spearman's rho is a valid test while Cronbach's alpha may not be the best statistic in this instance.

6.4.7.1 Cronbach's Alpha

From Appendix 32, Cronbach's Alpha was calculated using SPSS software.

6.4.7.2 Spearman's Rho

We considered using tau or rho since we are comparing sets of rankings. We therefore used Spearman's rho which is much easier to implement and is widely cited.

6.5 Development of a smartphone app for a genetics website, ALSoD

6.5.1 Searched online for direction

On Google Search Engine, I engaged in an extensive search using clauses like “How to convert website to mobile app”, “develop mobile app”, “wrap mobile website to app”. The first page containing 25 results were scanned through to see the websites which were most appropriate explaining how to develop either a mobile website or an app.

6.5.2 Choosing between a mobile website or an app

Initially, the idea of accessing the website in a more usable form on mobile devices was to develop an application (known as an ‘app’). Due to the similarity that exists between mobile websites and apps, some factors that determined our choice of a mobile website over an app are Target Audience, Budget Availability, Purpose Intended and Features Required [520]:

The team had to weigh the pros and cons of developing a website or an app or both. I decided to focus first on developing a mobile web-based platform because I wanted the content to work across all mobile platforms [512].

6.5.3 Optimization of webpage

Using our in-house built analytic data showing the most frequently viewed information coupled with the Google Analytics service configured for ALSoD website, the most frequently accessed webpages are on the pathogenicity of mutations, gene information and data analysis of mutation and patient data.

Google Analytics tool was configured in August 2012 to cross-check the results of visitors’ statistics built in-house. I based the data analysis on the three months (August, September and October). These first 3 months represented the period from when the Google Analytics tool was implemented to the point where I started the development of the mobile website (which will be referred to as the pre-mobile website period). The post-mobile website period is the second 3 months from November 2012 to January 2013 where the mobile website has been fully developed and an app implemented. In the pre-mobile website period, Pageviews (which is the total number of pages viewed) were analysed to discover the commonly visited pages on the website and counting repeated views of a single page.

The blog page was also added to allow for comments from researchers and patients who make use of the database extensively.

6.5.4 Design heuristics

Designing a mobile website that works on several platforms does not mean shrinking a complete webpage into a mini-size webpage. Users of mobile website have a pleasurable experience on a website when there are not too many tiny prints to read or they do not have to scroll left or right or they do not have to do too much typing. They are mostly interested in getting a quick answer by clicking once [512]. These points were utilized by designing a mobile website by [575]:

Creating a Separate Style Sheet.

Retaining some of the original images like the logo to prevent regular users of the full desktop website from being confused or creating an impression that they are working on a different website.

Reducing the size of images by a standard percentage.

Configuring the content layout of the screen to wrap texts to avoid any use of the horizontal bar.

Summarizing the information on the desktop version to fit a smaller screen.

6.5.5 Mobile Device Detection

On the Master page which allows us to create a consistent layout for the pages in our application, I first used the .NET framework mobile detection property (IsMobileDevice in the Request.Browser object). To detect if the request comes from a mobile device the .NET Framework fortunately provides the isMobileDevice property which returns a true value if the browser is a recognized mobile device or vice versa. This did not work fully on all mobile devices because some mobile device browsers disguise themselves as desktop browsers [576].

Detecting a mobile device was eventually determined here by the use of UserAgent strings sent by the browser from the mobile device to the server in conjuction with the isMobileDevice property as seen in Figure 123 and Figure 124.

```
<%@ Import Namespace="System.Web.UI.HtmlControls" %>
<%@ Import Namespace="System.Drawing.Imaging" %>
<%@ Import Namespace="System.IO" %>
<%@ Import Namespace="System.Configuration" %>
<%@ Import Namespace="System.Net" %>

<script language="C#" runat="server">

protected void Page_Load(object sender, EventArgs e)
    {
    DetectUserAgent();
    DataView dv1 = (DataView)SqlDataSource2.Select(DataSourceSelectArguments.Empty);
    Label1.Text = dv1.Table.Rows[0][0].ToString();
    }

    void DetectUserAgent()
    {
        string strUserAgent = Request.UserAgent.ToString().ToLower();
    if (strUserAgent != null)
        {
        if (Request.Browser.IsMobileDevice == true || ((strUserAgent.Contains("iphone") || strUserAgent.Contains("ipad") ||
        strUserAgent.Contains("blackberry") || strUserAgent.Contains("mobile") || strUserAgent.Contains("windows ce") ||
        strUserAgent.Contains("opera mini") || strUserAgent.Contains("palm") || strUserAgent.Contains("android") ||
        strUserAgent.Contains("samsung") || strUserAgent.Contains("nokia"))))
        {
            Response.Redirect("~/Mobile/index.aspx");
        }
        }
    }

</script>
```

Figure 123: A method for redirecting to the appropriate view

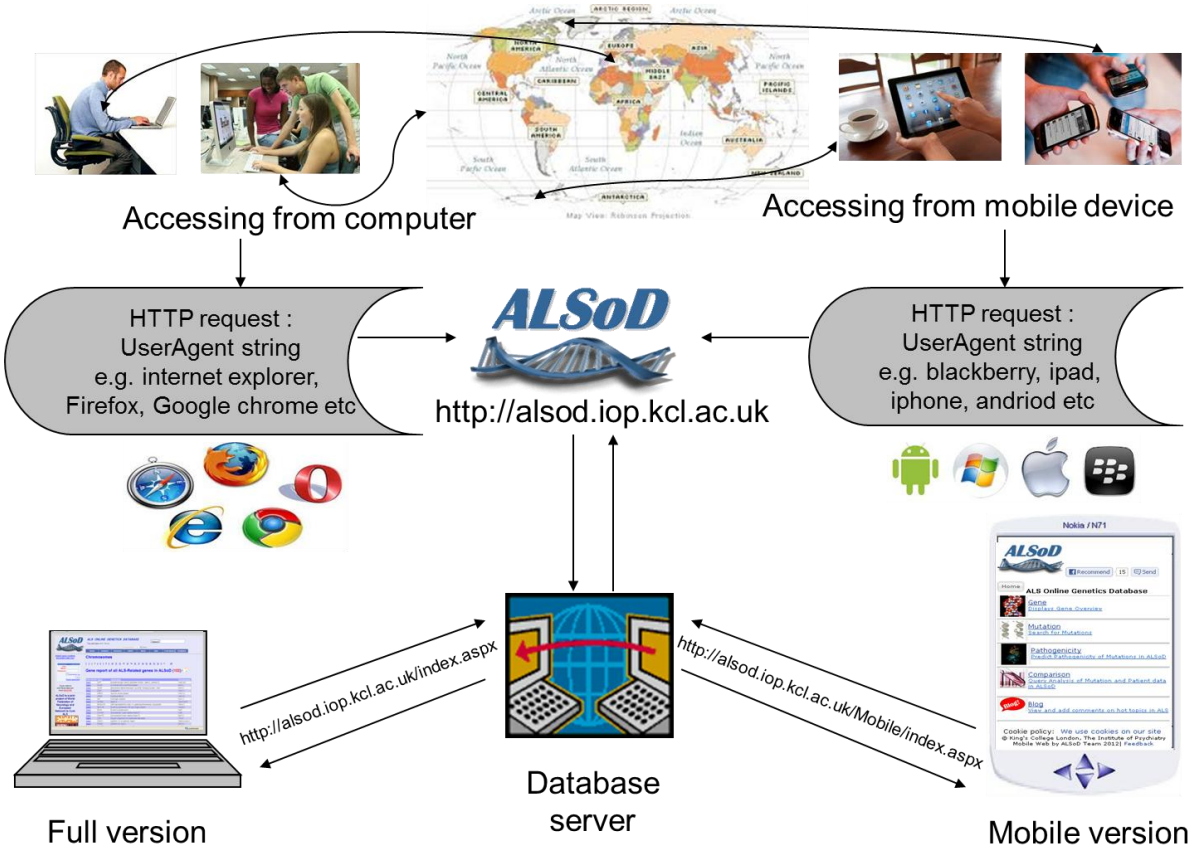


Figure 124: Overview of mobile website development

6.5.6 User Interface redirection

If the UserAgent string contains words like blackberry, palm, mobile, iphone, ipad (and other words used by mobile devices), then the user's device is redirected to http://alsod.iop.kcl.ac.uk/Mobile/index.aspx displaying a compact version of the website instead of the full version. That is,

If

[(IsMobileDevice = true) AND (UserAgent = true)] then (device = mobile)

View = "Full version"

Else

(device = computer)

View = "Mobile version"

6.5.7 Compact version of website

Web address link was sent to mobile phones of users through text messages and they were encouraged to click the facebook 'Recommend' button designed into the masterpage of the mobile version. Positive feedback were received from all - 2 Android users, 1 Nokia user, 5 Blackberry users, 2 iphone users and 3 ipad users. The success of the mobile version propelled us into developing an application.

6.5.8 Choosing to develop an app

Due to the intricacies involved in developing an app, I found out that the 5 most common smartphones to start developing applications for are iPhone, Blackberry, Android, Nokia98 and Windows Mobile. I decided to develop for Andriod since the Android market (now known as Google Play) allows an affordable fee for submission of apps unlike others. The market share of Android has grown and it is predicted as the most dominant mobile platform in the nearest future despite claims that lack of quality checks on regularly submitted sub-standard apps to the Android market is a downside [577].

With Android and Blackberry, the following processes were executed [578-582]: Install Eclipse software as the IDE; Download the Android SDK; Install the ADT plugin for Eclipse; BlackBerry Plug-in for Eclipse; Download the latest SDK tools and platforms using the SDK Manager; and Download Android and Blackberry Simulators.

The Plug-ins allows developers to develop, test and debug a Java application using the Eclipse IDE. The processes outlined above could be a daunting task for even a good programmer to get a grip of [583].

6.5.9 Using WebView to open a website

The process of developing an app easily and getting it into the App Market is to convert an already built mobile website into a native app. WebView object which is an in-app web browser is used to exhibit a website just as if the website is viewed on the browser of an Android smartphone [515]. Steps towards achieving this are [584]:

To build a new Android Application Project by navigating through File -> New-> Project -> Android -> Android Application Project -> type “ALSoD” as the project name, “android.uk.ac.iop.alsod” as the Package name and follow the rest of the wizard.

On the Activity Layout add a WebView by navigating through the Project Explorer to “res\layout\activity_main.xml” and open the XML file to add :

```
<android.webkit.WebView

android:id="@+id/wvwMain"

android:layout_width="fill_parent"

android:layout_height="fill_parent" />
```

On the manifest file add a permission code to ask the user for permission to access the internet before WebView can load a webpage. From the Project Explorer, open “AndroidManifest.xml” to add:

```
<uses-permission android:name="android.permission.INTERNET" />
```

On the Activity add some codes by navigating and opening the “ALSoD.java” file from the Project Explorer -> src -> com.projects.ALSoD -> MainActivity.java is a seen in Appendix 27.

6.5.10 Testing

When a regular website is developed, there is a necessity to view it on different browsers to ascertain compatibility across platforms and this is the same way a mobile application must be viewed on a wide diversity of mobile devices [515].

I initially downloaded emulators for iphone, andriod and nokia but a shortage of storage space on the computing system used made me to uninstall the software development kits. Since ALSoD has a facebook account and a Facebook 'Recommend' Button is embedded on the mobile masterpage, text messages were sent to as many contacts with internet access on my phone as possible. Blackberry Messenger was also sent to all my contacts asking them to view the web address (http://alsod.iop.kcl.ac.uk) on their phones and tablets. They were also asked to click on the facebook 'recommend' button so as to have an estimate of the number of users who were satisfied with the outcome of the display on their phone as seen in Figure 125.

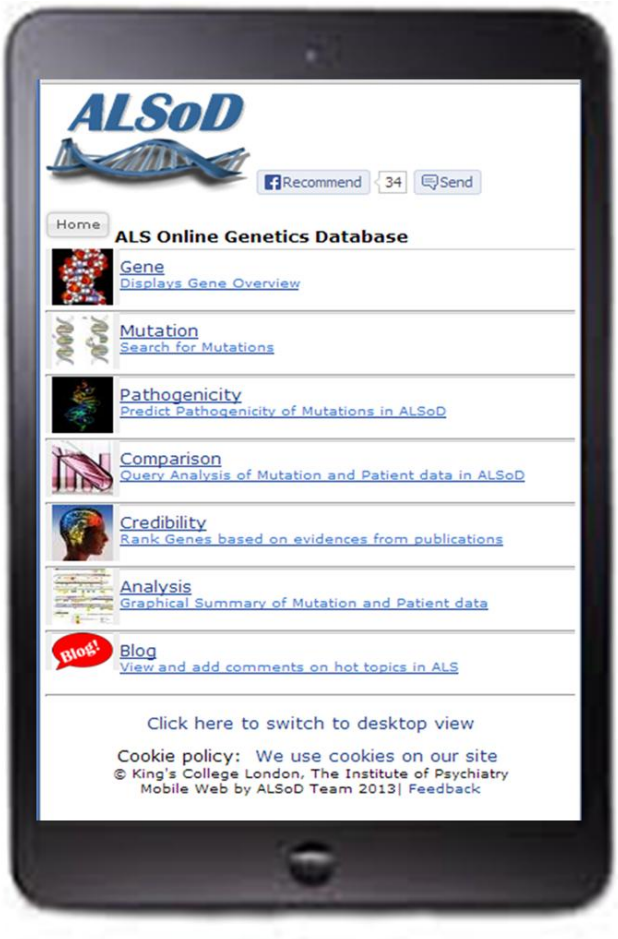


Figure 125: Mobile view of website optimize

Feedback from real users (2 Android users, 1 Nokia user, 5 Blackberry user, 1 iphone user, 3 ipad users) were collated in the development process.

Ammendments made after testing:

Caching for offline viewing [585-587]

Loading page when connecting

Make users aware of Cookies

Use option menu button [588] to display analysis webpages: interaction.aspx, credibility.aspx, analysis.aspx

6.5.11 App submission

From Eclipse, the application was compiled producing a .apk file. This file was submitted to a registered Google Play account with a generated keystore containing a private key [589]. The ALSoD app can be downloaded from Google Play and currently, our Google app account confirms that ALSoD app has had between 100 and 500 downloads as seen in Figure 126. The app then displays the ASP.NET website with no status bar or URL navigation on the screen.

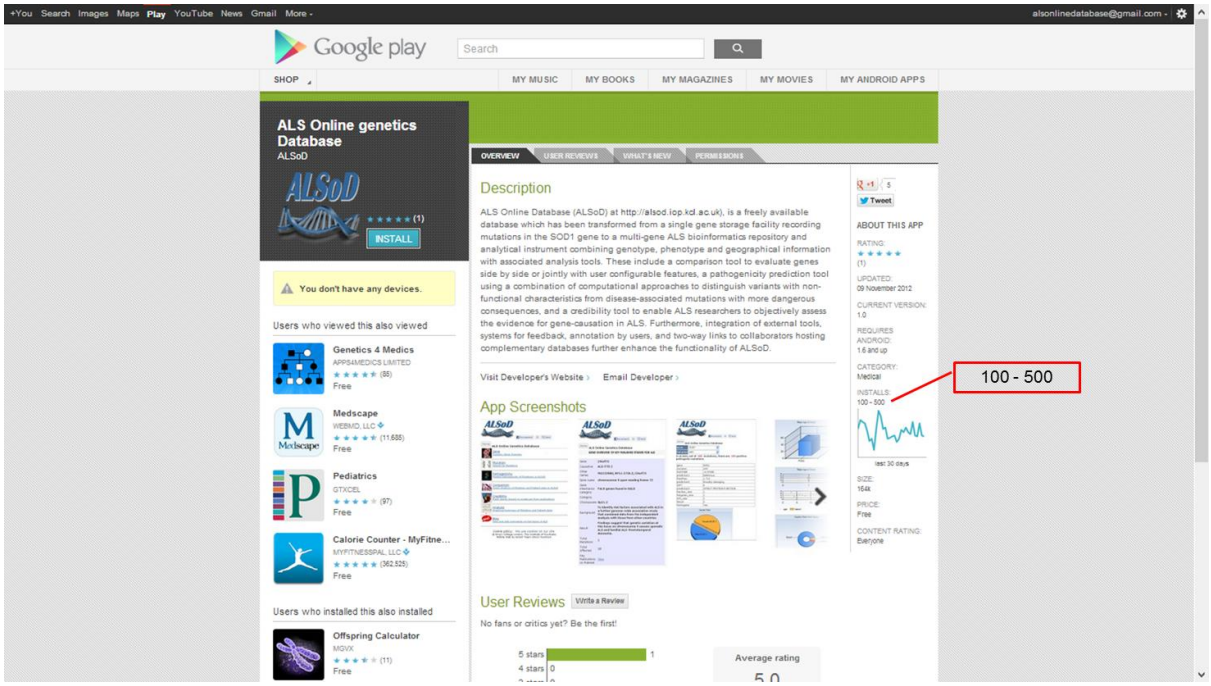


Figure 126: ALSoD Google play account showing number of downloads as at June 2013

6.5.12 Creating awareness

A marquee function scrolling a text from right to left was inserted on the desktop master page to create alertness for regular users of the website as seen in Figure 127. This was deliberately highlighted with a bright yellow bar for easy visibility. At symposiums and seminars, researchers were exposed to the recent development of the mobile app which has contributed to increased web traffic to ALSoD and a projected growth.

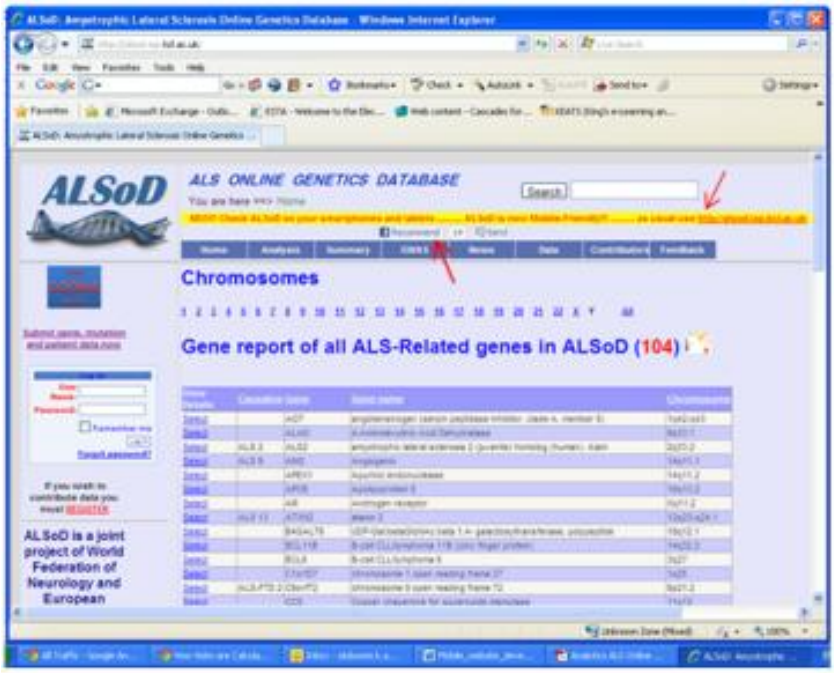


Figure 127: Desktop view creating awareness using a Facebook and marquee

6.5.13 Feedback from app users

During the various presentations of the mobile app development through posters and seminars, practical assessment of the website on mobile phones was carried out by attendants. Apart from the positive responses, questions asked, concerns raised and critical analysis given by the audience were recorded and considered.

6.5.14 Analysis of visits

The Google Analytics account for ALSoD was created in August 2012 to compare results generated by the two independent tools and to gain insight into the improvement of the design and content of the website. [590] The development of the mobile website started in October 2012 so I decided to compare visits from August 2012 to October 2012 and visits from November 2012 to January 2013.

Chapter 7 RESULTS

7.1 Keeping up with genetic discoveries in amyotrophic lateral sclerosis:

7.1.1 Data Submission

Once a user successfully logs into the ALSoD website for the purpose of submitting data, the user is able to submit a new gene or a new mutation or a replicated mutation data or patient data. A form is displayed on the website to fill in data which will eventually be submitted into the ALSoD SQL database. Submitted data are crosschecked regularly for any false submissions and the security of the database is regularly monitored. Below is a list of submissions and the tables where data are submitted to. Column list of each table is as shown in Appendix 24.

7.1.1.1 Submit new gene

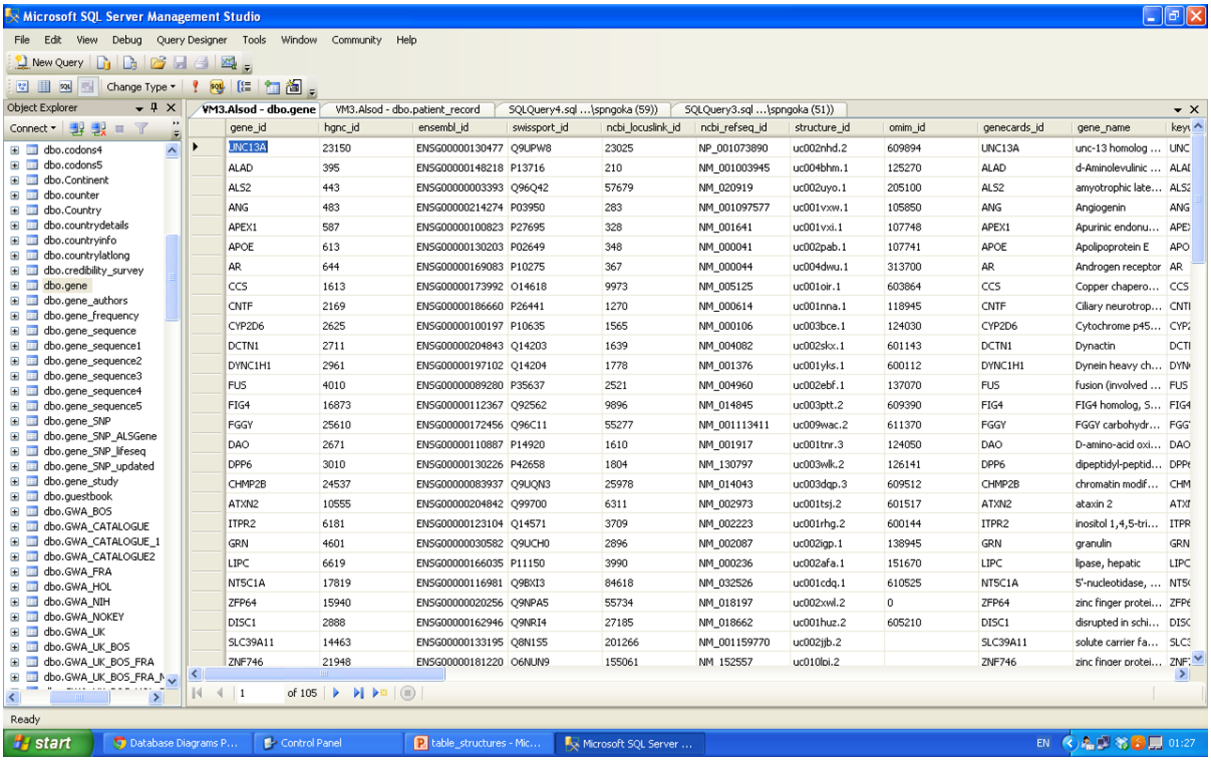


Figure 128: Submit new gene

7.1.1.2 Submit mutation data

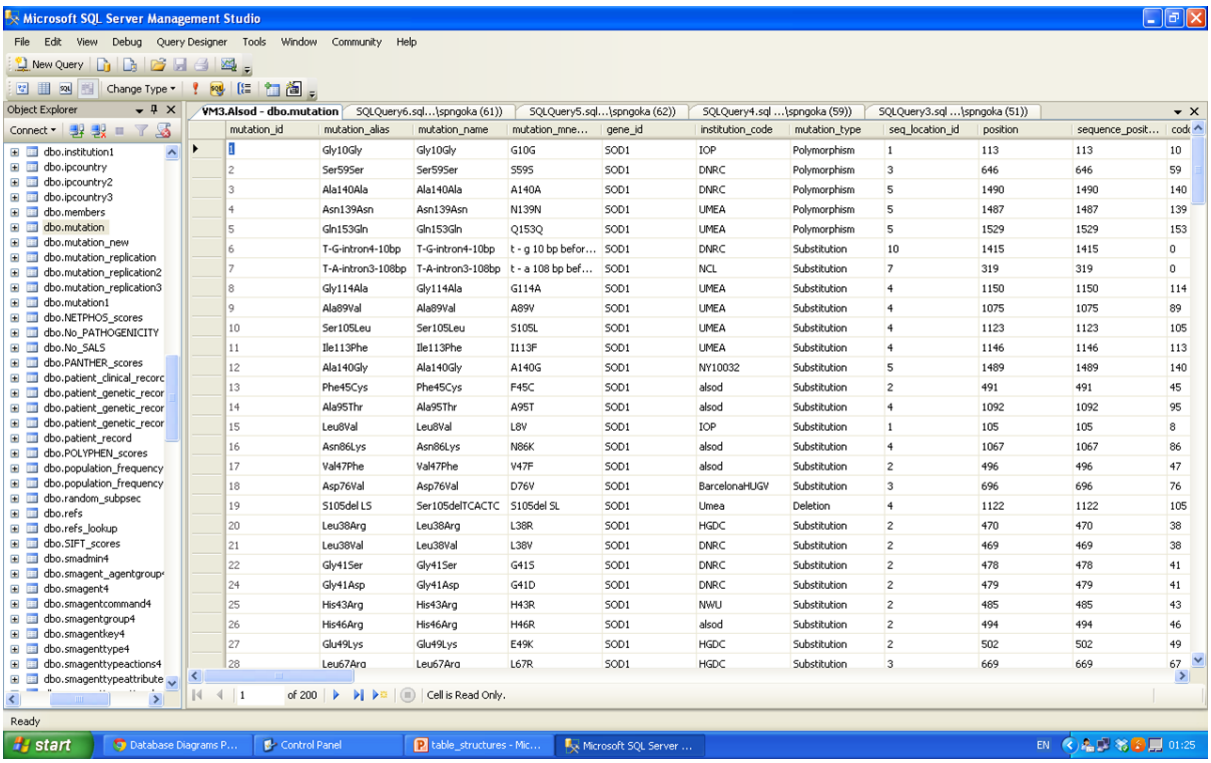


Figure 129: Submit mutation data

7.1.1.3 Submit patient data

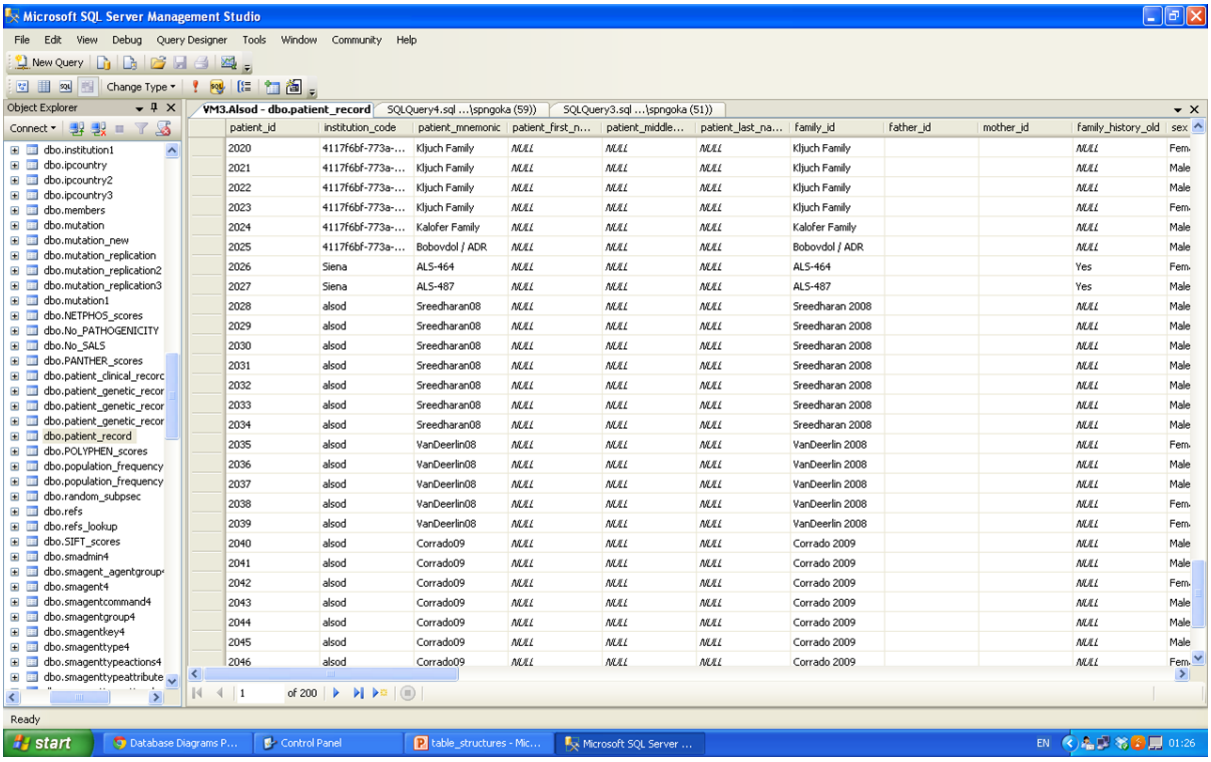


Figure 130: Submit patient data

7.1.1.4 Submit replicated mutation data

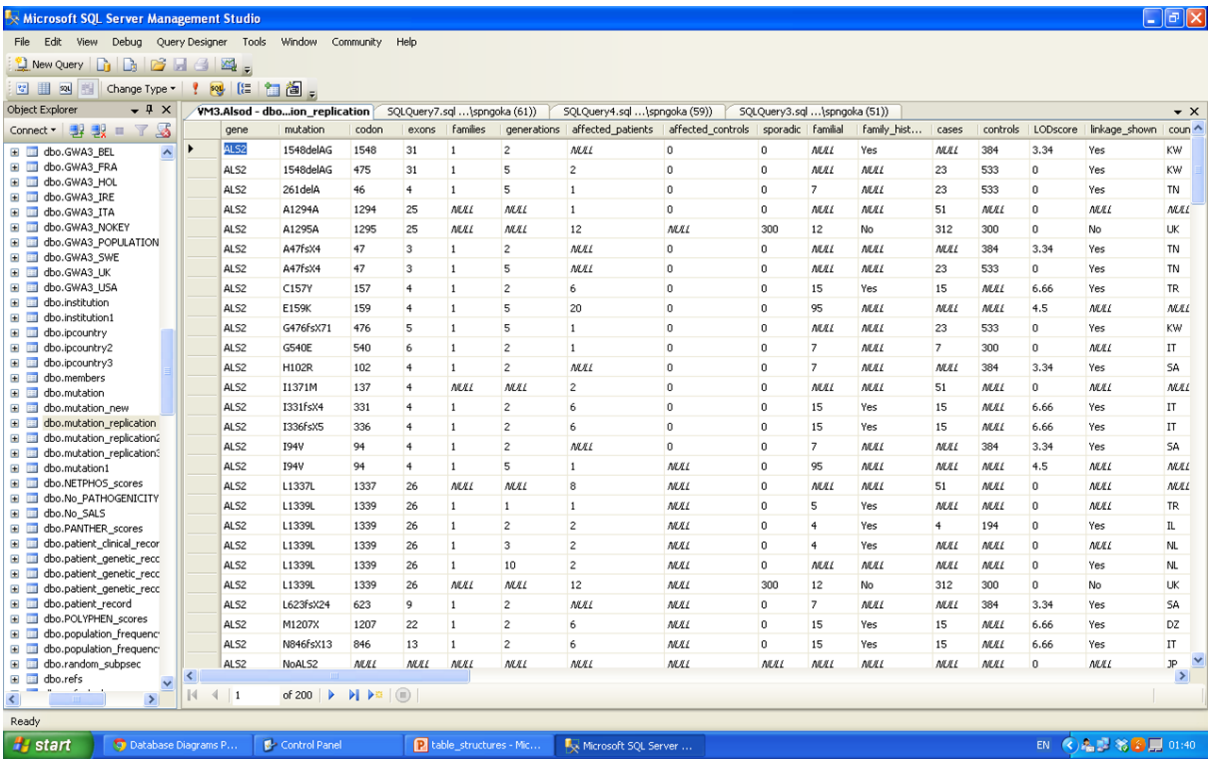


Figure 131: Submit replicated mutation data

7.1.2 Database Schema

The database schema now allows for flexibility and expansion because of changed table designs, rewritten queries, and implementation of stored procedures. Redundant tables have been removed, and a more supple structure is in place. The original database has been archived. New genes have been added to the tables, and a facility to easily add further genes designed. ALSoD now permits only registered users to submit novel gene, mutation, and patient data, and this is regularly validated by an ALS expert. Apart from system tables available on the database by default, there are 107 tables and 58 views on the database.

7.1.3 Search terms on search engines

A search term for novel mutations on a gene for example *SOD1* is displayed. As at September 2013, on Pubmed, 37 results were retrieved while on Google scholar, 330 results were retrieved (Figure 132).

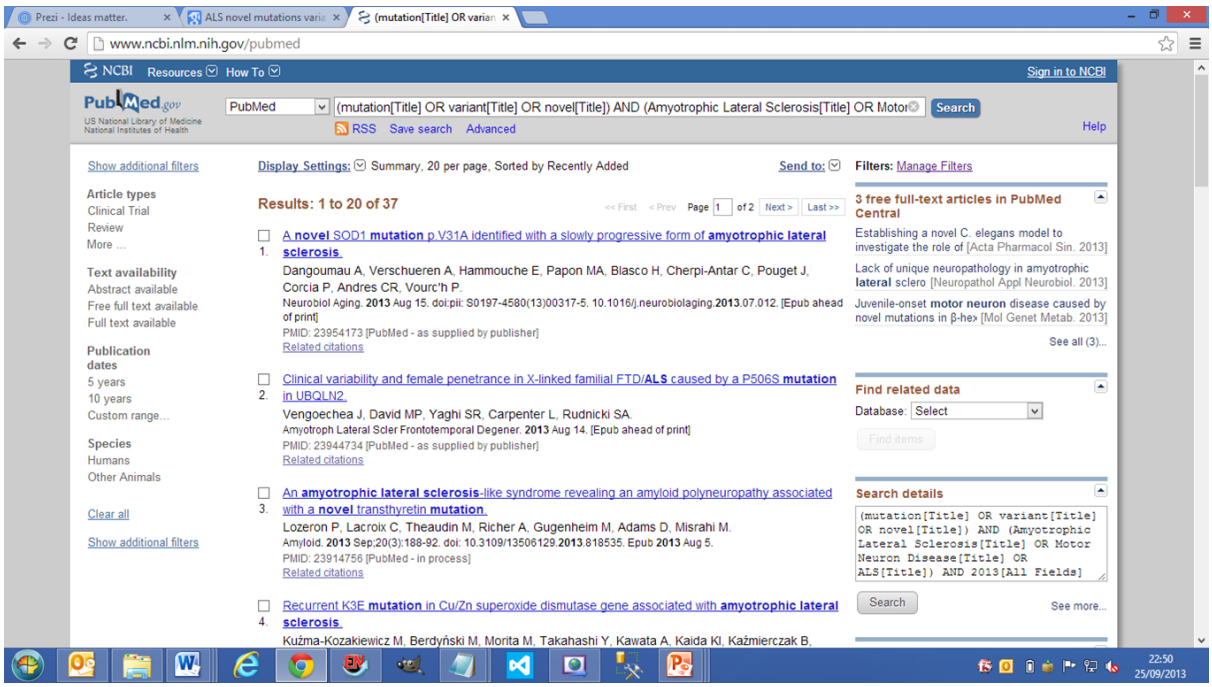


Figure 132: Screenshot of search terms on pubmed

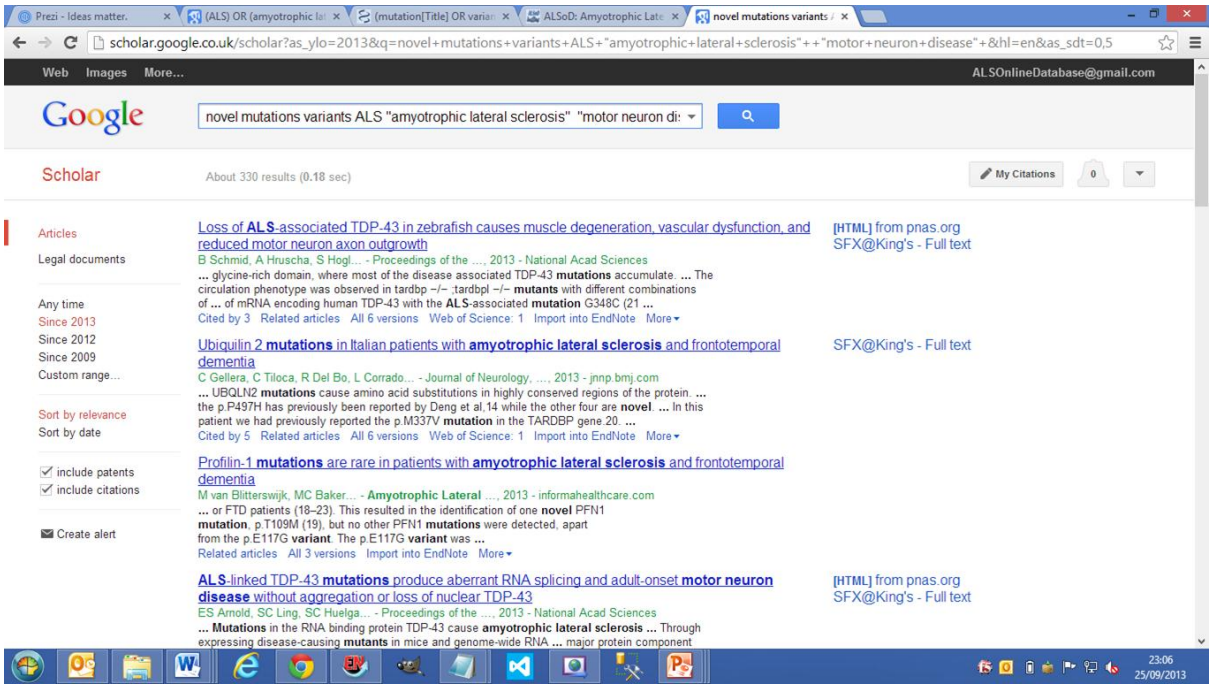


Figure 133: Screenshot of search terms on Google Scholar

Email Alerts, News page and List of last 10 submitted mutation and patient data are well designed and utilized to keep up with new genetic discoveries in the field of ALS. Initially, the website was designed to show the last 10 mutations submitted so, I modified the page to also show the last 10 patient data submitted and a graphical display of all mutation and patient data available in ALSoD in form of a bar chart.

A news page consisting of social media and blogs helps me to keep a tab on current issues in the ALS community and this is where news about genes and new technology for ALS is disclosed. This uses the Google RSS feed plugin to reveal current news. As at September 2013, the result is shown below:

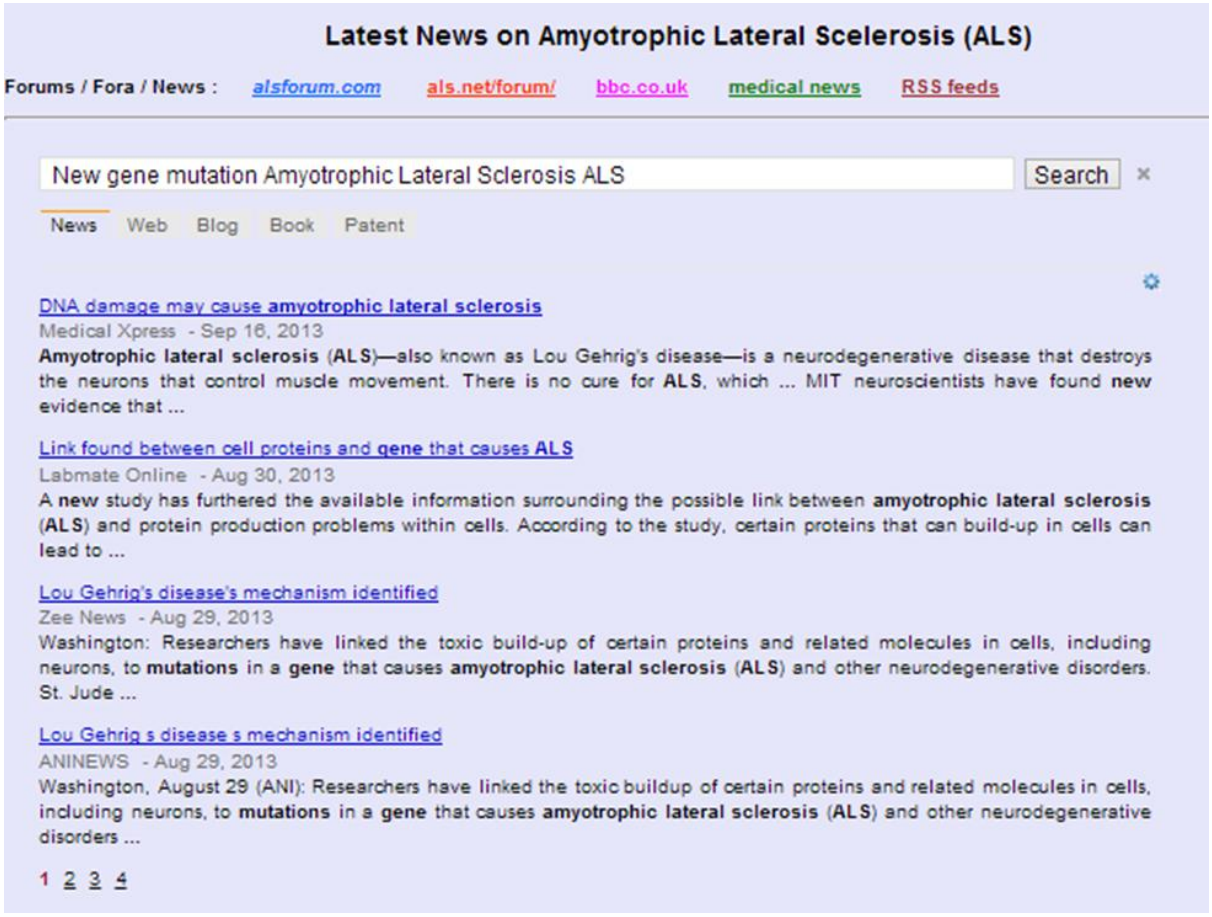


Figure 134: Latest News webpage

7.1.4 Contributors and Collaborators

There are currently 30 contributors from various international institutions. Due to regular updates on the website and constant involvement of experts in the field, contributors have increased from 12 to 30 since 2009. They voluntarily enter their details and a login and password are generated for them as regular submitters.

Andrew Martin's lab at UCL was the only collaborator with ALSoD before 2009 but the contact has since been lost due to changes in infrastructure and personnel change in his lab. Continuity has therefore been jeopardised and efforts made to restore the links from ALSoD has proved abortive.

So, at the moment, collaborations are with databases with freely available access to their resources online. Our current collaborators are geneMania group,

7.1.5 Unique Identifiers

Selecting a link automatically interrogates the relevant third party website with the appropriate genetic information. Thus, it is possible to identify the most up-to-date list of genes implicated in ALS, explore their interactions and pathways, and review the relevant literature.

7.1.6 Graphical display of phenotypic data

Customized and routine analyses of mutational relationships with phenotype are possible, either for genes with phenotypic data, such as SOD1, TARDBP, FUS, SETX and others, or as a general overview of familial and sporadic phenotypic patterns.

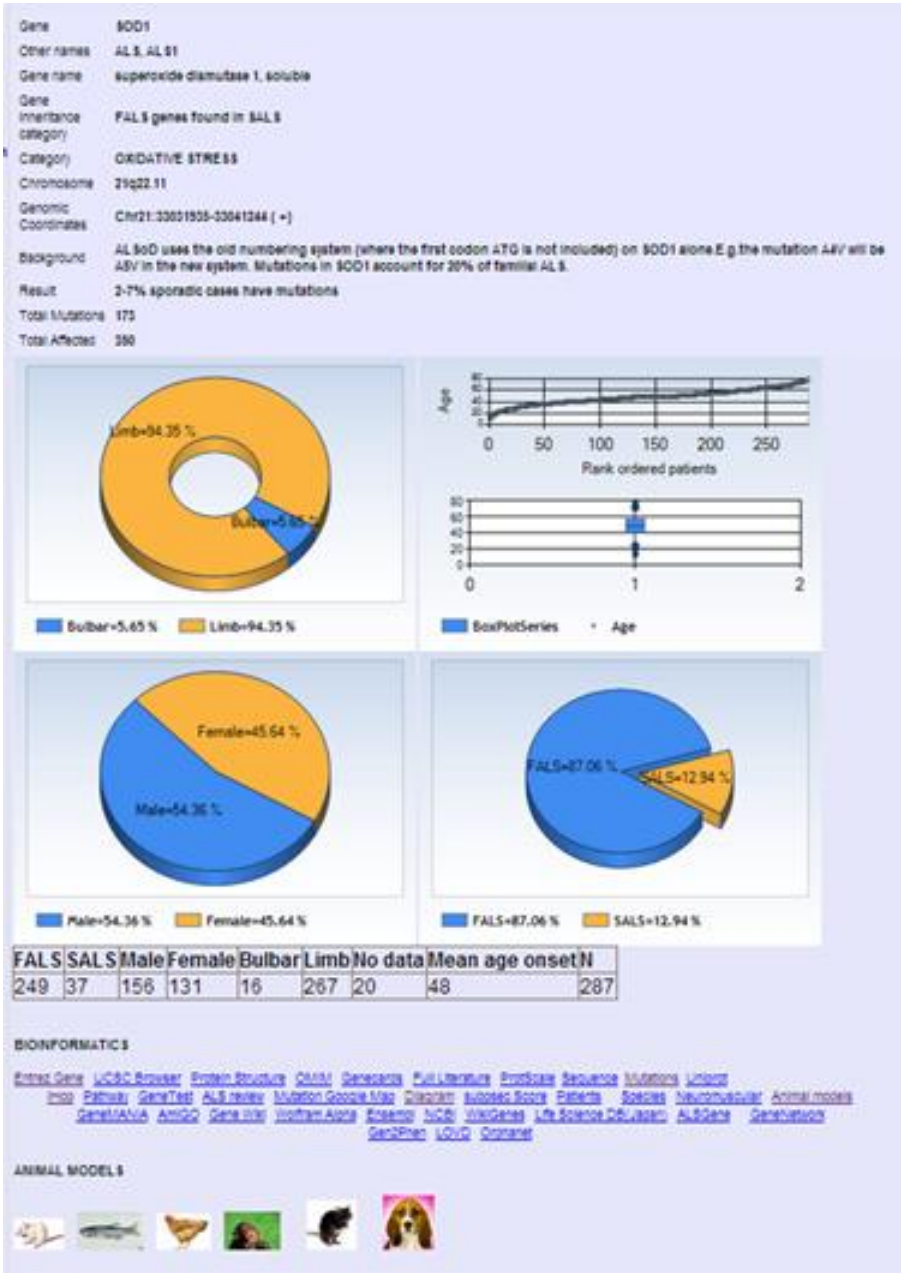


Figure 135: Gene Overview page

7.1.7 Mapping variants on Google Earth

The mapping of gene variants on Google Earth allows researchers or patients to view the geographical distribution of reported gene variations associated with ALS (Figure 3). The same computational method allows us to map the origin of users and show which countries predominantly access ALSoD.

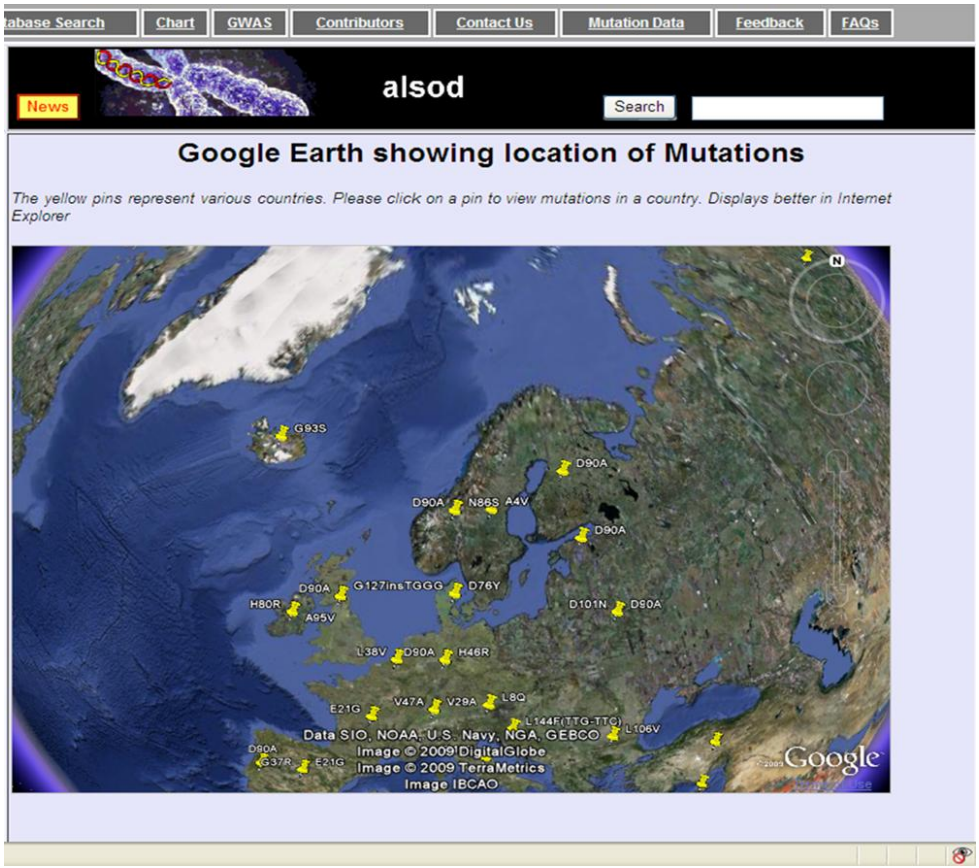


Figure 136: Mapping variants on Google Earth

7.1.8 Meta-analysis and on-the-fly analysis of association data

For users with their own association data, an on-the-fly analysis is available to combine the data available in ALSoD with unpublished user data that can be confidentially uploaded. The user data is formatted accordingly before upload and the result is fed back in minutes without storing users' data on the database.

7.2 ALSoD: A user-friendly online bioinformatics tool for Amyotrophic Lateral Sclerosis Genetics

7.2.1 Funding and Sponsorship

The ALSoD database is a joint project of the World Federation of Neurology and European Network for the Cure of ALS, and is funded through grants from the ALS Association, Motor Neurone Disease Association. The research leading to these results has received funding from the European Community's Health Seventh Framework Programme FP7/2007-2013 under grant agreement number 259867. Others are ALS Canada, ALS Therapy Alliance, and MNDA Iceland.

7.2.2 Programming

Open source programming software such as JavaScript, C#, T-SQL, Perl, XML, and VB.NET integrated under the ASP.NET platform are implemented to write codes and scripts. ALSoD uses the Microsoft .NET framework. Microsoft SQL server 2008 is used to manage the database stored on the VM3 server of the Institute of Psychiatry, King's College London. Microsoft Visual Web Developer 2008 Express Edition is used to develop the user-interface dynamic Web pages. Google AJAX search and Google Earth API have been combined to overlay geographical mutation information of ALS-related genes on the globe.

7.2.3 Web Design

A facelift was given to the Web page and some pages were redesigned for better visual representation of data. The Graphical User Interface allows data to be interpreted and viewed at a glance instead of using the tabular format of viewing data. A summary of ALSoD's site map is shown in Figure 137 below using the treeview. A full version as shown on http:alsod.iop.kcl.ac.uk/maps/treeview.aspx is seen in Appendix 13.

Figure 137: Sitemap of ALSoD

7.2.3.1 MasterPage

The Master page which is a file on "../master/MasterPage.master" allows the web designer to create a regular look and behaviour for all the pages (or group of pages) on the website. It outlines placeholders for the content, which can be superseded by content pages. The final outcome is an amalgamation of the master page (template) and the content page (content to display).

7.2.3.2 Homepage

The publications are directly accessed using the first author linked to pubmed, year, title and the external web link to the full paper (where full text is publicly available online). Key publications with a reported novel mutation or novel gene or novel patient data are displayed. The database was restructured to enhance

flexibility. ALSoD version 3 began with the manual curation of more genes and mutations from review publications [209, 508] while patient data were curated from publications referenced in ALSoD available in Pubmed.

7.2.3.3 Analysis webpage

For example, a user wishes to predict gene interactions between the two genes (TARDBP and FUS). These two genes are checked on the list, a list of the selected genes are displayed for confirmation and once the user is satisfied, the EXECUTE link is clicked to link with the webservice on GeneMANIA website. The output is then displayed on the same ALSoD webpage without having to move back and forth between websites. The RESET link can be clicked if another query is desired.

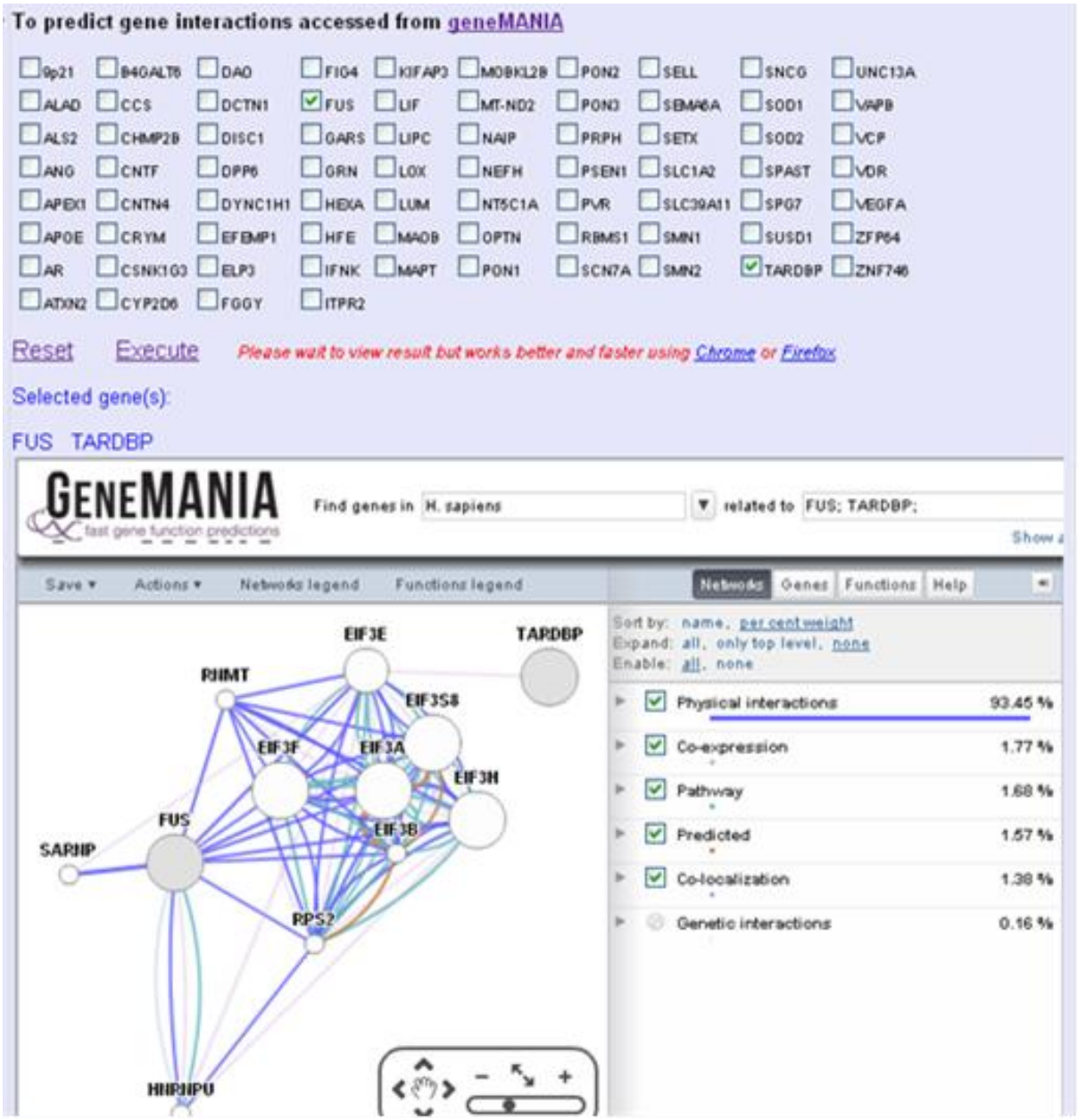


Figure 138: Interaction result using GeneMANIA

To compare genetic data between any two selected genes, the dropdownlist aids the user to view the list of genes available for analysis on the database. For example, a user wants to compare the genetic data of all patients between two genes (FUS and TARDBP). Click on the first dropbox to select FUS and click on the second dropbox to select TARDBP. Leave the default age range of 0 to 100 for all patients, then click on the COMPARE button. A graphical representation showing a box plot on Age of Onset and 3 pie charts on Site of Onset (Limb/Bulbar ratio), Gender proportion (Male/Female ratio) and Inheritance type (FALS/SALS ratio) are displayed side by side for the two selected genes. On the same page, a joint analysis of both genes and a tabular list of the patient and mutation data with references to the publications where these data are curated are also displayed.

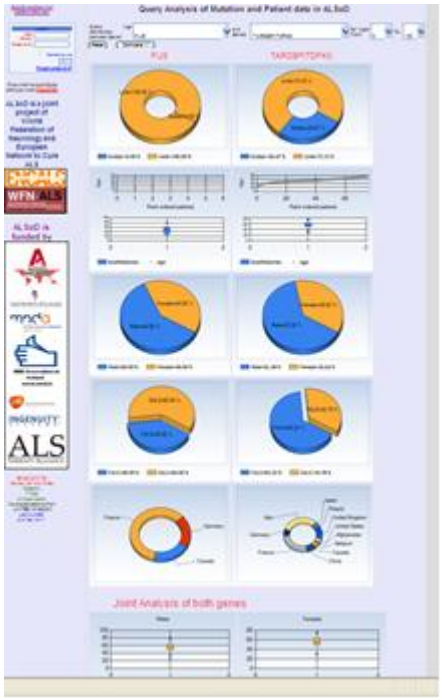


Figure 139: Side by side comparison result between FUS and TDP43

For detailed analysis, a list of genes with mutations are shown in checkboxes. A more detailed analysis of selected genes with a range of age of onset with or without further selections (on the site of onset, gender and family history) are performed and graphically displayed as seen in Figure 140.

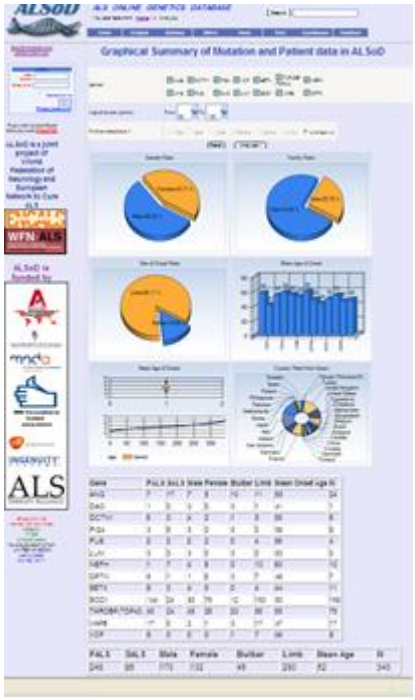


Figure 140: Detailed analysis of a selected gene

7.2.3.4 Summary webpage

Report page displays a graphical statistical representation of the mutations (Mean, Median, Variance, Standard deviation) of all genes available in ALSoD. These analyses are not produced from static data and any change or update in the dataset will recalculate and be displayed online dynamically. Images below show the patient summary report on each gene, the top 20 most frequent mutations in patients and mutation data summary with their gene effects deposited in ALSoD respectively.

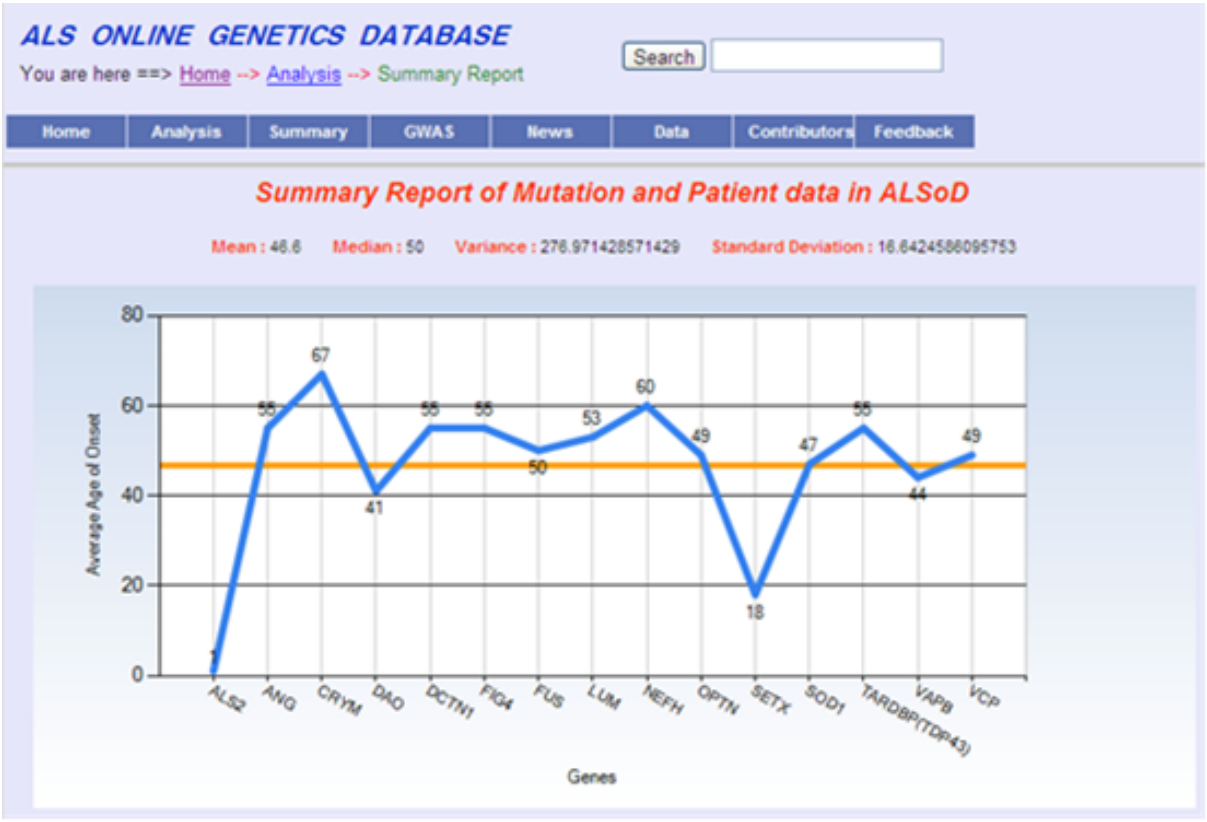


Figure 141: Summary report of genetic data

7.2.3.5 GWAS webpage

7.2.3.5.1 Whole Genome

A user is required to log on with a registered username and password in order to display a static haploview graphical overview of all submitted genome wide results in ALS and a dynamically generated graph with selected threshold. These two graphs help to verify the same information graphically but the clickable graph navigates to the NCBI Single Nucleotide Polymorphism website for more bioinformatics information on the selected SNP (Fig 1). Despite selecting a default threshold of 3, the page is loaded in 18 seconds due to the large volume of data. Although the ten SNPs shown in Table 3 have unknown clinical associations, with larger sets of data analyzed in future there is a possibility of finding significant SNPs with known clinical associations.

If all the 5 populations (UK, Boston, Holland, France, NIH) are selected, it takes approximately 20 seconds to display the result as shown.

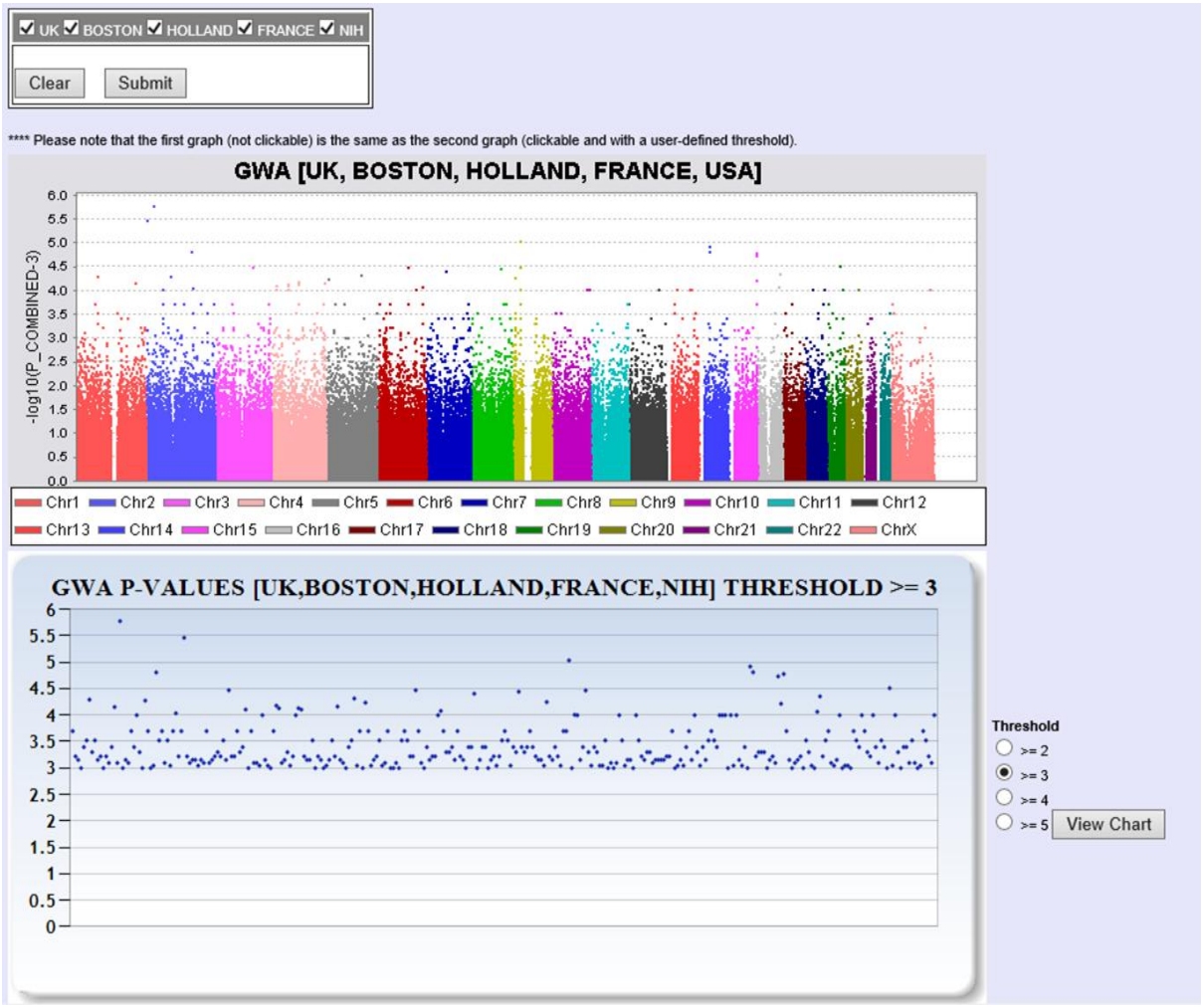


Figure 142: Displaying GWA result of 5 populations on Whole Genome

The top result is a screen dump of the analysis manually performed on haploview while the second result is a dynamically generated scatter chart corresponding with the screen dump as a form of check.

7.2.3.5.2 On-the-fly meta analysis

The on-the-fly tool enables a registered user to upload and analyse their data by comparing these data with the dataset in ALSOD. Data must be in a text format with four (4) compulsory fields: Chromosome, SNP, BP and P-Value fields respectively. There must be no headings on the first row, each field must be separated with a TAB and each row terminated with a RETURN key. An on-the-fly graphical representation of this comparison is shown and the significance of the SNPs are shown and linked to a bioinformatics source on NCBI webpage for further information. This takes an average of 65 seconds to appear on the page.

7.2.3.5.3 Verifying a particular SNP

This shows the p values from each source of dataset. The combined p value is derived after logging on with a username and password. This is to conceal the privacy or identity of the DNA information as

suggested in the study which proves that the probability that a person or relative participated in a GWA study can be assessed. It is a known fact that much effort is put into making experimental data publicly available in order to combine the data with other studies [12] The populations to be combined are selected using a set of checkboxes which later produces a static haploview graph, a dynamically generated graph with selected threshold and a graph showing significant SNPs of the selected chromosome as seen in Figure 143.

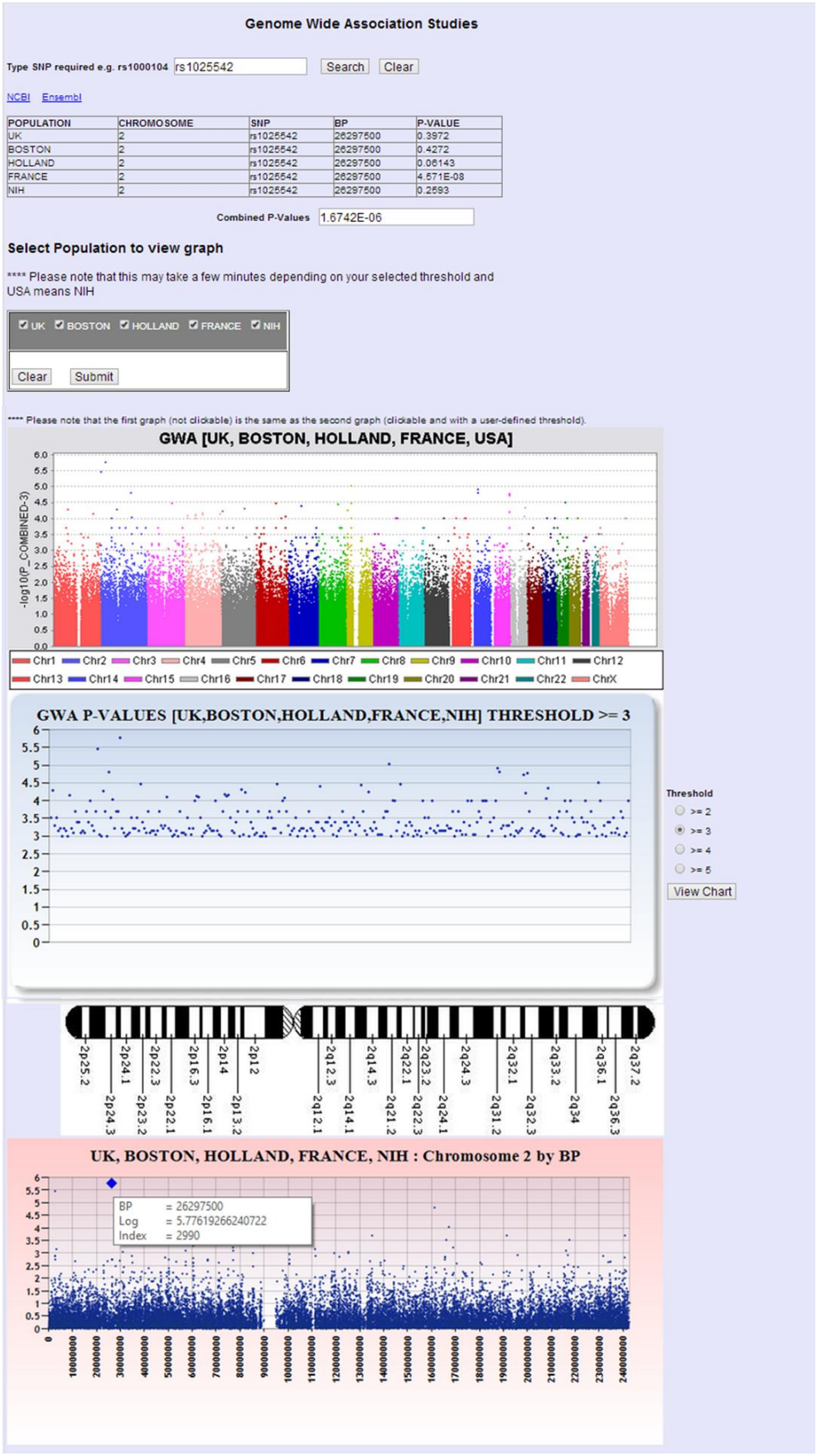


Figure 143: Displaying output of querying database for a single SNP

7.2.3.5.4 Chromosomal view on UCSC genome browser

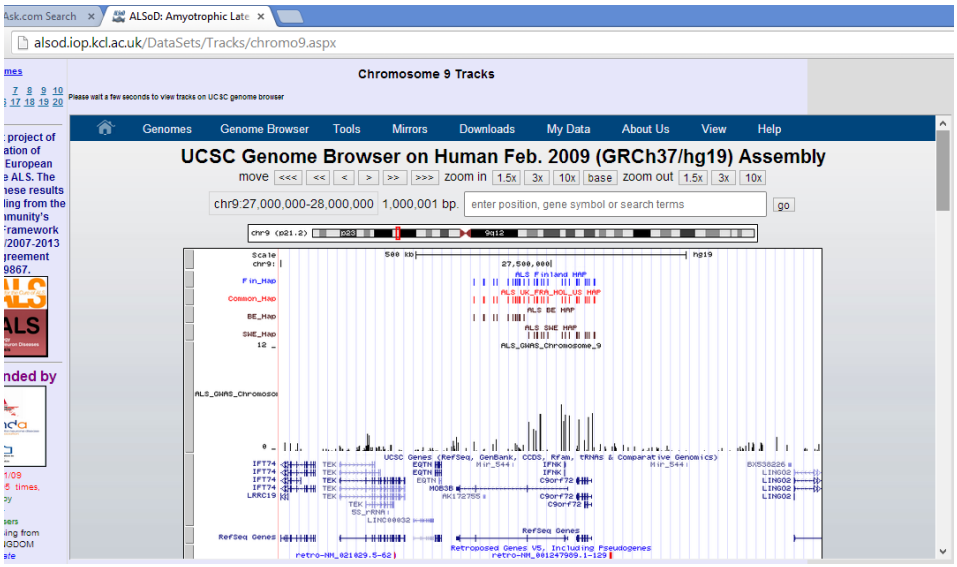


Figure 144: UCSC Genome Browser on chromosome 9

GWAS by Fogh 2013 is a summarizes SNPs statistic table from logistic regression analyses combined in the GWAS meta-analytical study which includes 8 independent studies with 13,225 individuals analyzed for 6,138,740 overlapping markers [544].

Genome Wide Association Study by Fogh et al.

A genome-wide association meta-analysis identifies a novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. Fogh I, et al. Hum Mol Genet. 2013 Nov 20.

Summary SNPs statistic table from logistic regression analyses combined in the GWAS meta-analytical study using the software METAL www.sph.umich.edu/csg/abecasis. The meta-analysis includes eight independent studies shown in the table below with 13,225 individuals analyzed for 6,138,740 overlapping markers.

Type SNP required e.g. rs17763230 rs34517613 Search Clear

CHROMOSOME	SNP	BP	PVALUE
17	rs34517613	23634379	1.105E-08

Sample size analyzed in the ALS-GWAS meta-analysis

Source	Sample Ancestry	Sample size (N)	Cases (N)	Control (N)	Genotyping Illumina Arrays	SNPs before imputation	SNPs after imputation
SLAGEN Consortium	Italian	3959	1982	1977	660K	657366	8391895
UMC Utrecht	Dutch	911	461	450	317K	317503	8328981
Utrecht, Umeå , Leuven	Dutch, Swedish, Belgian	2806	1364	1442	370K	317503	8372645
MGH, King's College , Evry	Northern American, British, French	3916	1710	2206	370K	307790	8370626
Beaumont Hospital, Dublin	Irish	432	221	211	550K	561466	8268635
NIH publicly available, Coriell biobank	Northern American	547	276	271	550K	555352	8325506
MND DNA bank and RADIANT study	British	2252	663	1589	610K	471994	8372429
NIH publicly available	Italian	747	500	247	550K	500002	8312642

Figure 145: GWAS by Fogh 2013

7.2.3.6 News webpage

Using the Google search engine integrated into ALSoD, external links to published articles on the latest news in the ALS research community are displayed using the keywords “New gene mutation Amyotrophic Lateral Sclerosis ALS”. The current information are separated into 5 categories: News, Web, Blog, Book and Patent.

Other useful hyperlinks such as ALS forum, ALS.NET forum, BBC website news, Medical news and RSS feeds from pipelines are also available on the webpage for users.

7.2.3.7 Data webpage

This page is popularly accessed by researchers due to the ability to download mutation data or patient data or 1000 genome or EVS genome in a spreadsheet format. It is a very useful page for cross-checking data on the database for the purpose of transparency. The genome data displayed on a flat file on the webpage shows global count and frequency, counts and frequencies from different populations like South America, Asia, Africa and Europe.

7.2.3.8 Contributors webpage

Collaboration with other databases like ALSGene, ALS mutation database , HGMD, Uniprot, FALS Connect, PANTHER, SIFT, PolyPhen, PubMed, Google scholar and Jalview has improved the functionality of ALSoD. Integration with these external but free databases have been made possible with unique identifiers for each gene.

7.2.3.9 Feedback webpage

Main users of the database are Researchers, Clinicians, Patients and their families. So, feedback is essential to the development of the database to direct our improvement design and tools.

7.2.4 On-the-fly Meta-Analysis of Genome Wide Association Studies in ALS

Published data from our laboratory and Landers’ group are available in ALSoD [403, 455]. An on-the-fly meta-analysis allows users to upload their formatted data (according to specifications available on the help page) confidentially to examine the joint analysis result. A simplified list of published GWAS data in ALS alone derived programmatically from the very long list on the GWAS Catalogue [591] is displayed for easy use by researchers without having to download and write codes to filter the gigantic dataset themselves. Information tabulated are rs number (linked to corresponding details on NCBI Single Nucleotide Polymorphism webpage), chromosomal region, gene, first author (linked to the pubmed article), year, initial sample size, replication sample size, risk allele, risk allele frequency, P value, odd ratio, confidence interval

and platform used. A total number of SNPs with a link to other comprehensive GWAS databases like HuGE Navigator [592], HGVbaseG2P [477] and ALSGene [539] are shown. ALSoD only gives users basic up-to-date GWAS information as a more comprehensive and unbiased GWAS meta-analysis can be queried directly from the ALSGene website or a hyperlink from a gene overview page [486].

This GWAS analytical tool is displayed in three ways:

7.2.4.1 Verifying a particular SNP

This shows the p values from each source of dataset. The combined p value is derived after logging on with a username and password. This is to conceal the privacy or identity of the DNA information as suggested in the study which proves that the probability that a person or relative participated in a GWA study can be assessed. It is a known fact that much effort is put into making experimental data publicly available in order to combine the data with other studies [524]. The populations to be combined are selected using a set of checkboxes which later produces a static haploview graph, a dynamically generated graph with selected threshold and a graph showing significant SNPs of the selected chromosome as seen in Figure 143.

7.2.4.2 Whole genome

A user is required to log on with a registered username and password in order to display a static haploview graphical overview of all submitted genome wide results in ALS and a dynamically generated graph with selected threshold. These two graphs help to verify the same information graphically but the clickable graph navigates to the NCBI Single Nucleotide Polymorphism website for more bioinformatics information on the selected SNP (Figure 142). Despite selecting a default threshold of 3, the page is loaded in 18 seconds due to the large volume of data. Although the ten SNPs shown in Table 3 have unknown clinical associations, with larger sets of data analyzed in future there is a possibility of finding significant SNPs with known clinical associations.

7.2.4.3 On-the-fly analysis

The on-the-fly tool enables a registered user to upload and analyse their data by comparing these data with the dataset in ALSOD. Data must be in a text format with four (4) compulsory fields: Chromosome, SNP, BP and P-Value fields respectively. There must be no headings on the first row, each field must be separated with a TAB and each row terminated with a RETURN key. An on-the-fly graphical representation of this comparison is shown and the significance of the SNPs are shown and linked to a bioinformatics

source on NCBI webpage for further information. This takes an average of 65 seconds to appear on the page.

7.2.5 Collaborations and Embedded Tools

ALSoD uses third party open source bioinformatics tools to embed computational analysis within the database using Java applets. For example, in Figure 146, a screenshot of the Multiple Alignment and Mutations on SOD1 gene using a combination of Claustalw and Jalview [Waterhouse et al., 2009] is used to provide multiple sequence alignments in other species for selected genes. GeneMANIA [Warde-Farley et al., 2010] allows users to select genes of interest for prediction of interactions. A Google Earth API is used for viewing maps of mutation, risk, and exposure distributions. Because many ALS gene variants are found in both familial and apparently sporadic ALS, a two-way link out to the ALSGene database provides evidence of association to complement the genotype–phenotype correlation available from familial ALS information in ALSoD [Lill et al., 2011]. A similar link out to fALS Connect, which is a collaboration between multiple interested agencies in the United States, including the patient organization The ALS Association and the research group The Northeast ALS (NEALS) Clinical Trials Consortium, makes ALSoD relevant for patients and carers as well as the scientific community. The database is adopted into the Human Variome Project (http://www.humanvariomeproject.org) and the GWAS Phenomap Project (http://www.gwascentral.org/gwasphenomap).

For example, if SOD1 gene is selected, the webpage links to the webpage http//:alsod.iop.kcl.ac.uk/Applets/SOD1species.aspx. A screenshot of the Multiple Alignment and Mutations view of the SOD1 gene using a combination of Claustalw and Jalview is displayed.

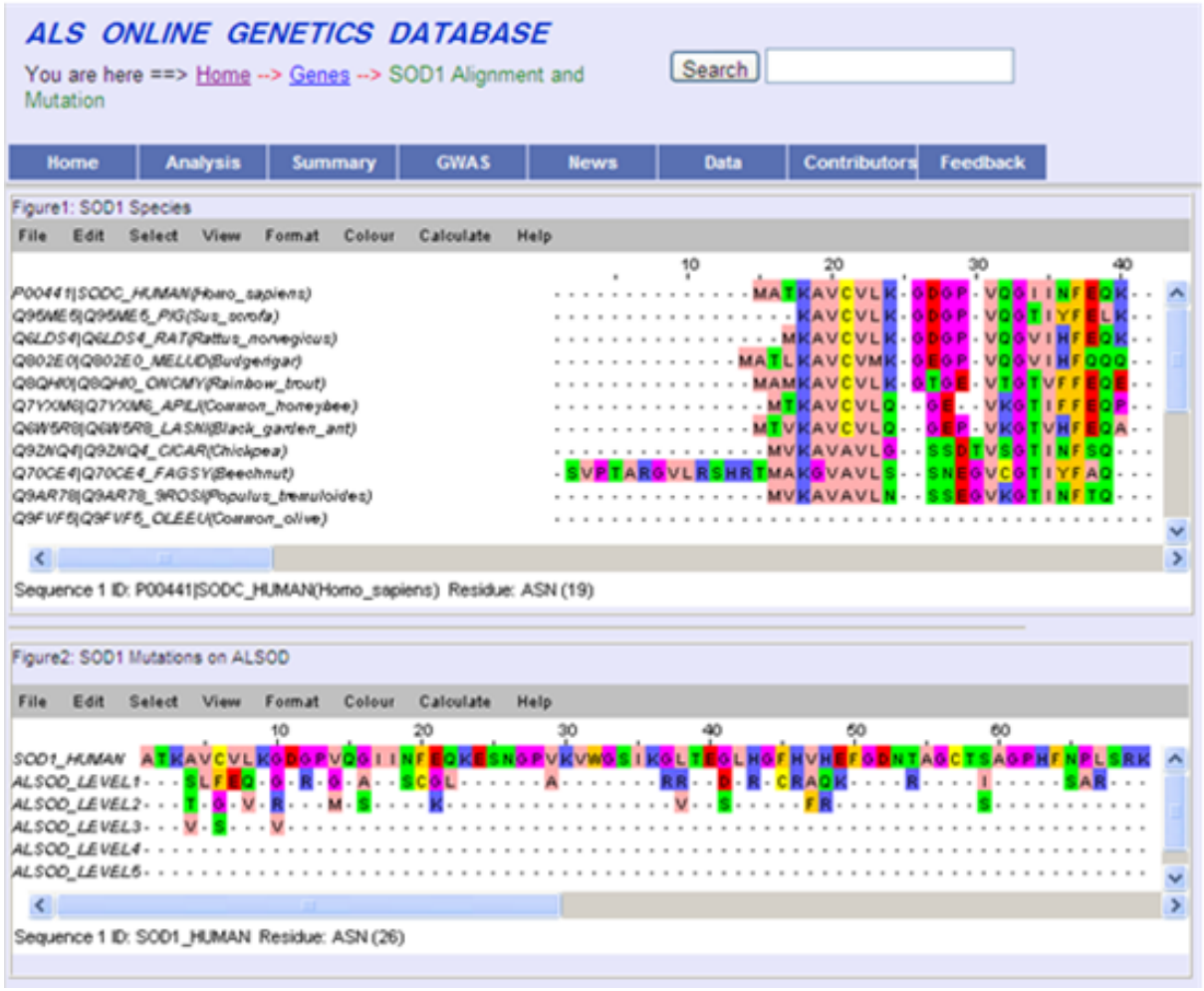


Figure 146: Multiple Alignment on SOD1

7.2.6 Integrated Bioinformatics Links

To avoid bias, users can retrieve gene-specific information through external links which have been programmed automatically for each gene, and which open in new windows. Unique identifiers are utilized by systematically linking to broad databases and bioinformatics tools freely available online. The scientific and nonscientific external links integrated into ALSoD include HGNC [White et al., 1997], Entrez Gene [Maglott et al.], UCSC Browser [Fujita et al.], Protein Structure [Rose et al.], OMIM [Amberger et al., 2011; Amberger et al., 2009], Genecards [Safran et al.], ProtScale [Gasteiger et al., 2005], KEGG [Kanehisa et al., 2000], Uniprot [Jain et al., 2009], iHop [Hoffmann and Valencia, 2004], Pathway in KEGG [Kanehisa et al., 2000], GeneTest [Pagon et al., 2002], AmiGO [Carbon et al., 2009], Ensembl [Hubbard et al., 2009], NCBI [Sherry et al., 2001], Life Science DB (Japan) [Yoshida et al., 2010], ALSGene [Lill et al., 2010], GeneWiki [Huss III et al., 2008], WolframAlpha (Maret), and WikiGenes [Hoffmann, 2008].

7.2.7 Feedback

Feedback is gained in two main ways: a Facebook page for ALSoD http://www.facebook.com/srch.php#!/pages/ALSoD/307667685943735, and a direct feedback page on the ALSoD Website. Comments are publicly displayed and a reCAPTCHA tool displays texts readable only by human users to prevent spammers from infiltrating the system (Appendix 42). A news page generates automated summaries of ALS genetics news; and surveys conducted through the freely available online survey tool “SurveyMonkey” are embedded in the user interface.

A blog page was created to allow users make comments and a social media page consisting of Twitter and Facebook was developed to allow those not connected to these social media have access to information shared by others. This form of communication helps to answer the kind of question posted by a user saying, “What is the latest mortality rate of ALS?”.

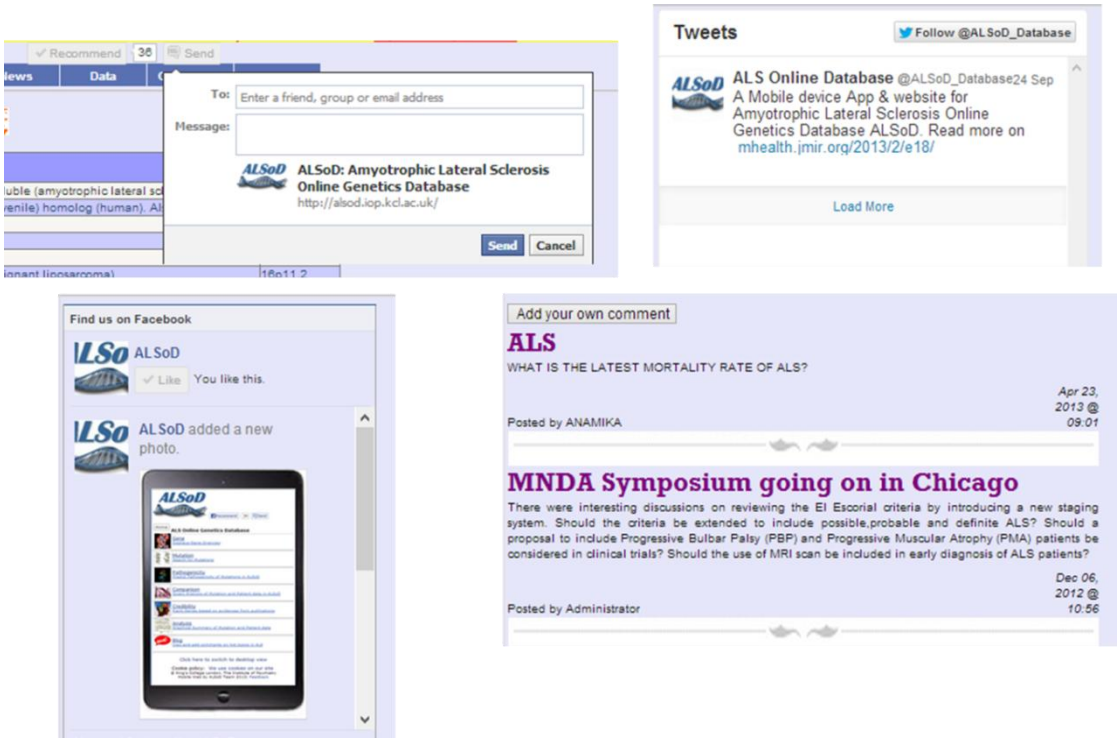


Figure 147: Social media accounts of ALSoD

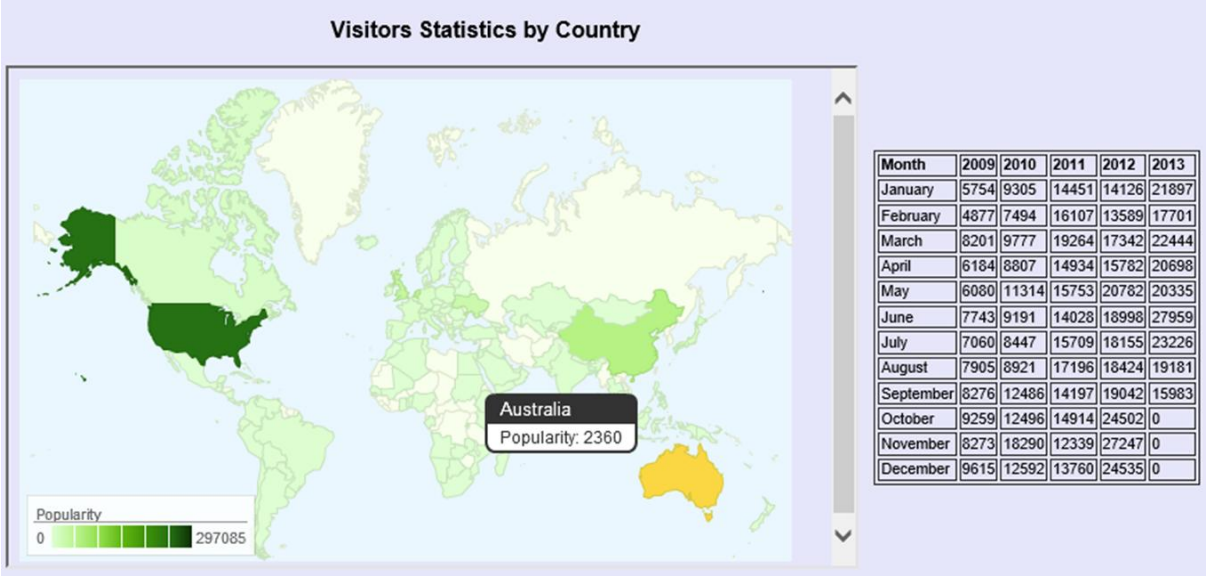
7.2.8 Tracking Visitors

In addition, by tracking the registered country of origin of page viewing and download requests, accessibility of the ALSoD database to the international ALS community can be monitored directly. A tracking tool was developed to monitor the influx of visitors to the website by identifying unique ipaddresses per day. This revealed that there were visitors from over 125 countries with 18,290 users in the month of November 2010 alone (Figure 4) which requires an enormous amount of storage facility on the database.

This has helped to compare visits between 2009 and 2010 as shown in Figure 5. An example of a day's visit to the ALS Online database is seen in Figure 6

7.2.8.1 Visitors' webpage

From Feedback -> Visitors' statistics which is on http://alsod.iop.kcl.ac.uk/charts/index.aspx. A graph of visits from January 2009 to August 2013 shown below confirms that ALSoD is constantly and increasingly being accessed by users. For example, in Australia, so far, there have been 2360 visits.



Month	2009	2010	2011	2012	2013
January	5754	9305	14451	14126	21897
February	4877	7494	16107	13589	17701
March	8201	9777	19264	17342	22444
April	6184	8807	14934	15782	20698
May	6080	11314	15753	20782	20335
June	7743	9191	14028	18998	27959
July	7060	8447	15709	18155	23226
August	7905	8921	17196	18424	19181
September	8276	12486	14197	19042	15983
October	9259	12496	14914	24502	0
November	8273	18290	12339	27247	0
December	9615	12592	13760	24535	0

Figure 148: Visitors' statistics

Every year, there is a higher number of visits than the previous year even though it dwindles in the number of visits monthly due to downtimes or holidays across the globe.

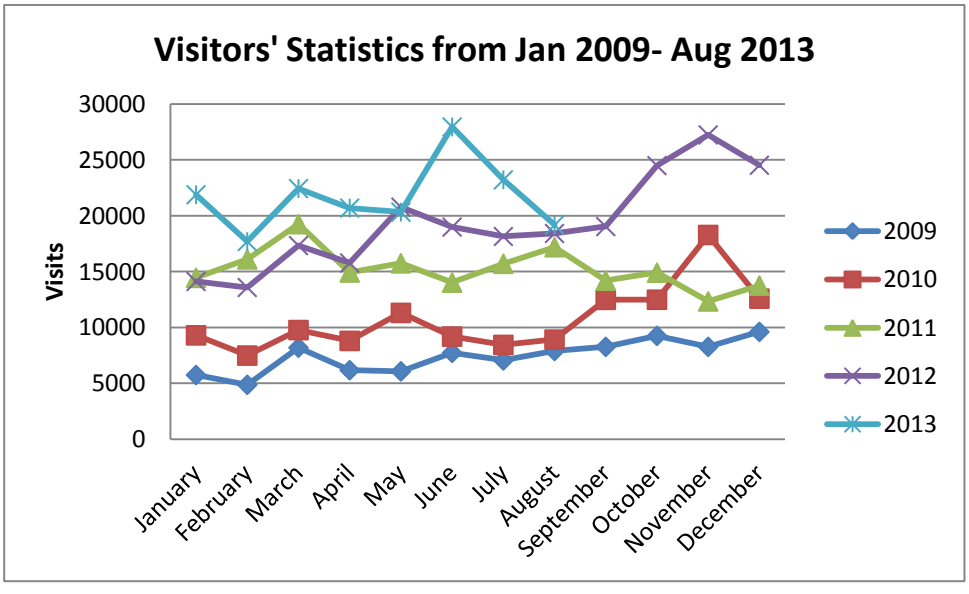


Figure 149: Graph of Visitors' statistics

7.2.8.2 Google Analytics tool

As a form of parallel monitoring of the use of ALSoD worldwide, I used the Google analytics plugin tool to find out information on how frequently ALSoD is accessed, from where, the popular webpages users visit etc. Just within a year, from August 2012 to present date in September 2013, about 6550 unique visitors have visited the website from around the world.

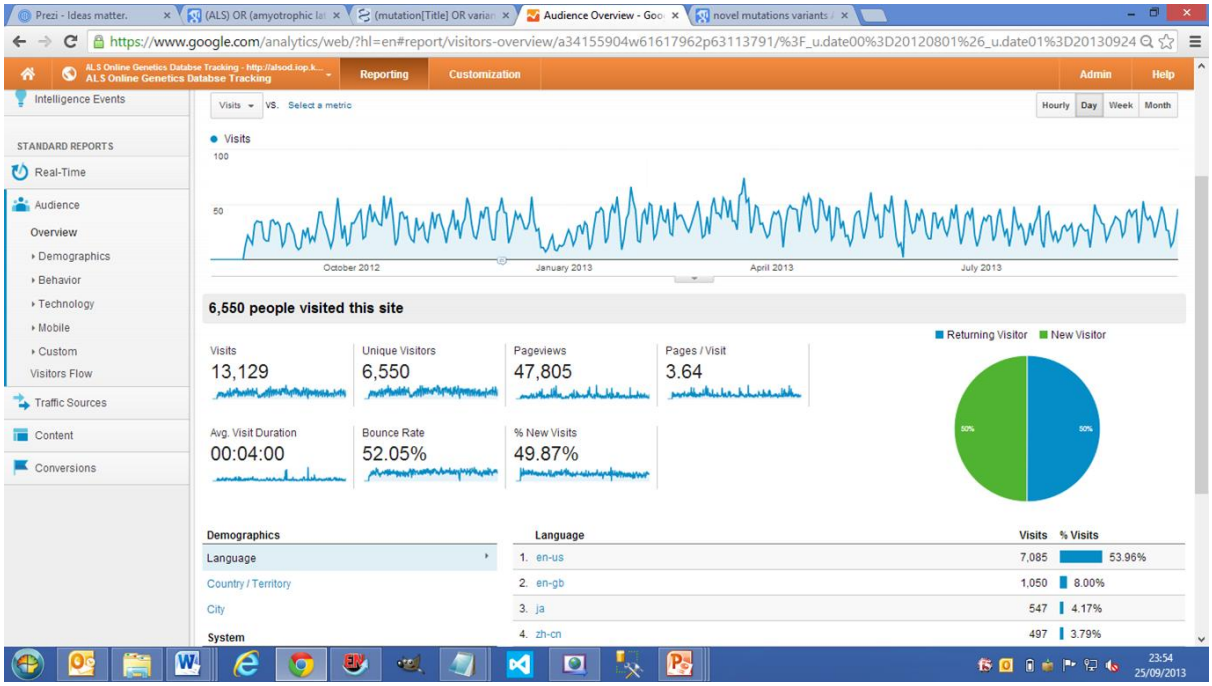


Figure 150: Google Analytics display

7.2.8.3 ALSoD Impact

Apart from the number of unique users and the number of times the database has been visited as shown on the homepage, other interesting statistics are found on http://alsod.iop.kcl.ac.uk/charts/index.aspx.

As at the 18th of October 2013, visitors' statistics from http://alsod.iop.kcl.ac.uk/charts/index.aspx are: Total number of countries where ALSoD has been accessed (152); Highest access from USA, China, Netherlands and UK; and Total number of visitors daily (like previous day 17th October) is 706 (varies from day to day).

From Google scholar, using the search term "alsod.iop.kcl.ac.uk", a total of 234 results were displayed on "http://scholar.google.co.uk/scholar?start=15&q=%22alsod.iop.kcl.ac.uk%22&hl=en&as_sdt=0,5" while some publications cited the ALSoD database using the old web address. So with the search term "alsod.org", a total of 174 results were shown on "http://scholar.google.co.uk/scholar?hl=en&q=%22alsod.org%22&btnG=&as_sdt=1%2C5&as_sdtp=".

7.2.9 ChangeLog

The ALS Online Genetics Database website was initially located at http://www.alsod.org and is now located at http://alsod.iop.kcl.ac.uk. This change in web address and hosting of this valuable tool became imminent due to an insatiable crave for huge storage. ALSoD allows users to submit new gene, mutation and patient data. It also has various tables in the database schema, storing data on genes, mutations, patients, codons, gene sequence, trinucleotides, country details, users and stored procedures. It uses the Microsoft .NET framework. Microsoft SQL server 2008 is used to manage the database stored on the VM3 server of the Institute of Psychiatry, KCL (size about 10830 MB). Microsoft Visual Web Developer 2008 Express Edition is used to develop web pages (current size about 707MB containing 387 files and 53 folders).

More than 110 ALS-related genes have now been added to the database with a current total of 431 mutations (195 pathogenic) and 589 patient data. 15 of the mutations are unpublished except in ALSoD. ALSoD webpages have been visited over 280,000 times since 2009 by more than 22,900 unique visitors from 140 countries. There are 26 registered contributors excluding those from the host institution. 33 different publications have cited functionalities or updates available on ALSoD.

7.2.10 Search keyword(s) externally or internally

Using Google Search engine for both searches, the external search finds a keyword on the World Wide Web while the internal search finds a keyword on ALSoD website only. For example, using the keyword : ‘C9orf72’

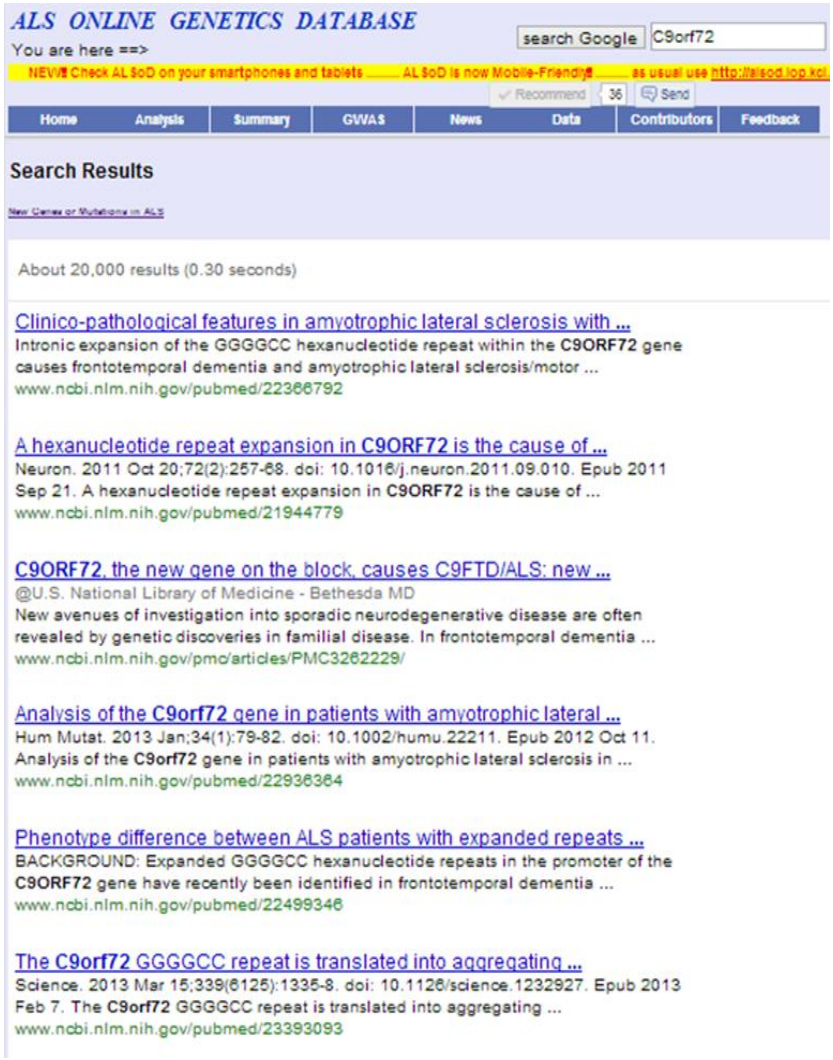


Figure 151: Search output using Google search engine interface for external search

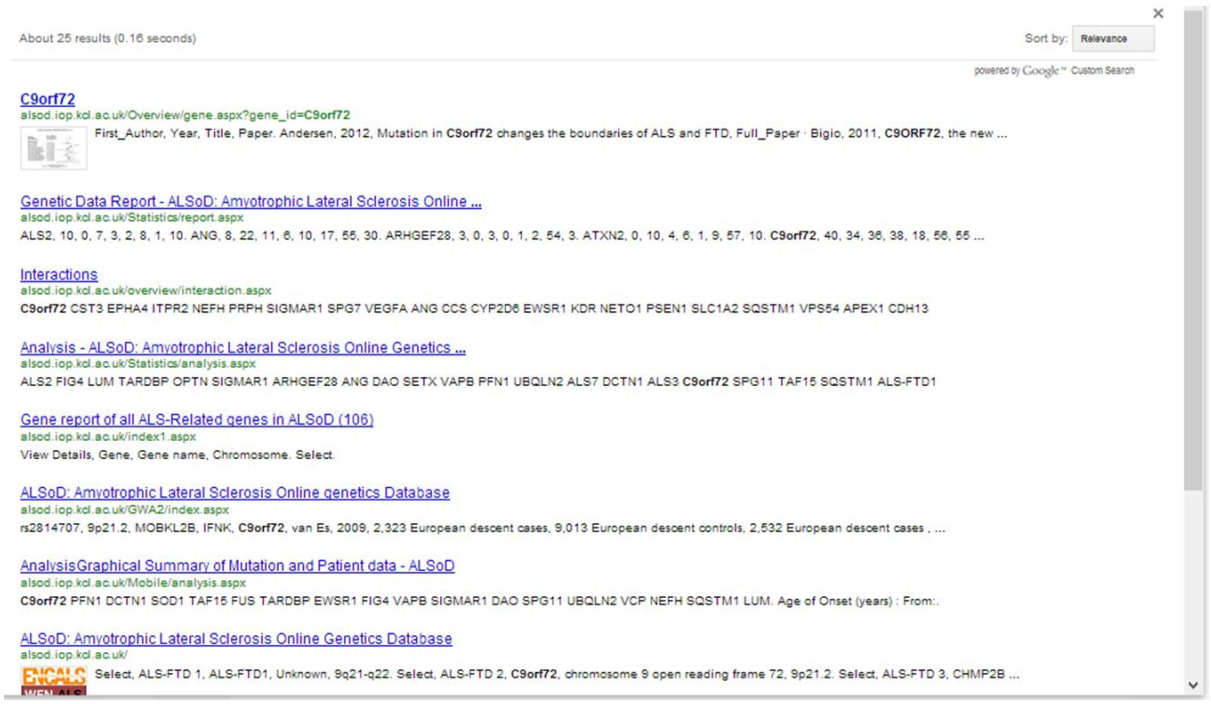


Figure 152: Search output using Google search engine interface for internal search

7.3 Credibility analysis of putative disease-causing genes using bioinformatics

PRISMA revision [593] with respect to development and reporting of results were taken into consideration as seen in Appendix 8.

7.3.1 Data collection

Genes with at least one publication suggesting involvement in adult onset familial ALS were studied [108]. I excluded genes with limited clinical data, absent mutational data or unreplicated results. Publicly listed variants for the included genes derived from ALSGene, Uniprot, ALS Mutation and HGMD databases were merged with variant lists in ALSoD, and filtered for duplicates (Appendix 8).

7.3.2 Pathogenicity analysis using bioinformatics tools

Database is queried as discussed in 6.4.2.6 dispaying a summary of predicted pathogenicity shown in Figure 153.

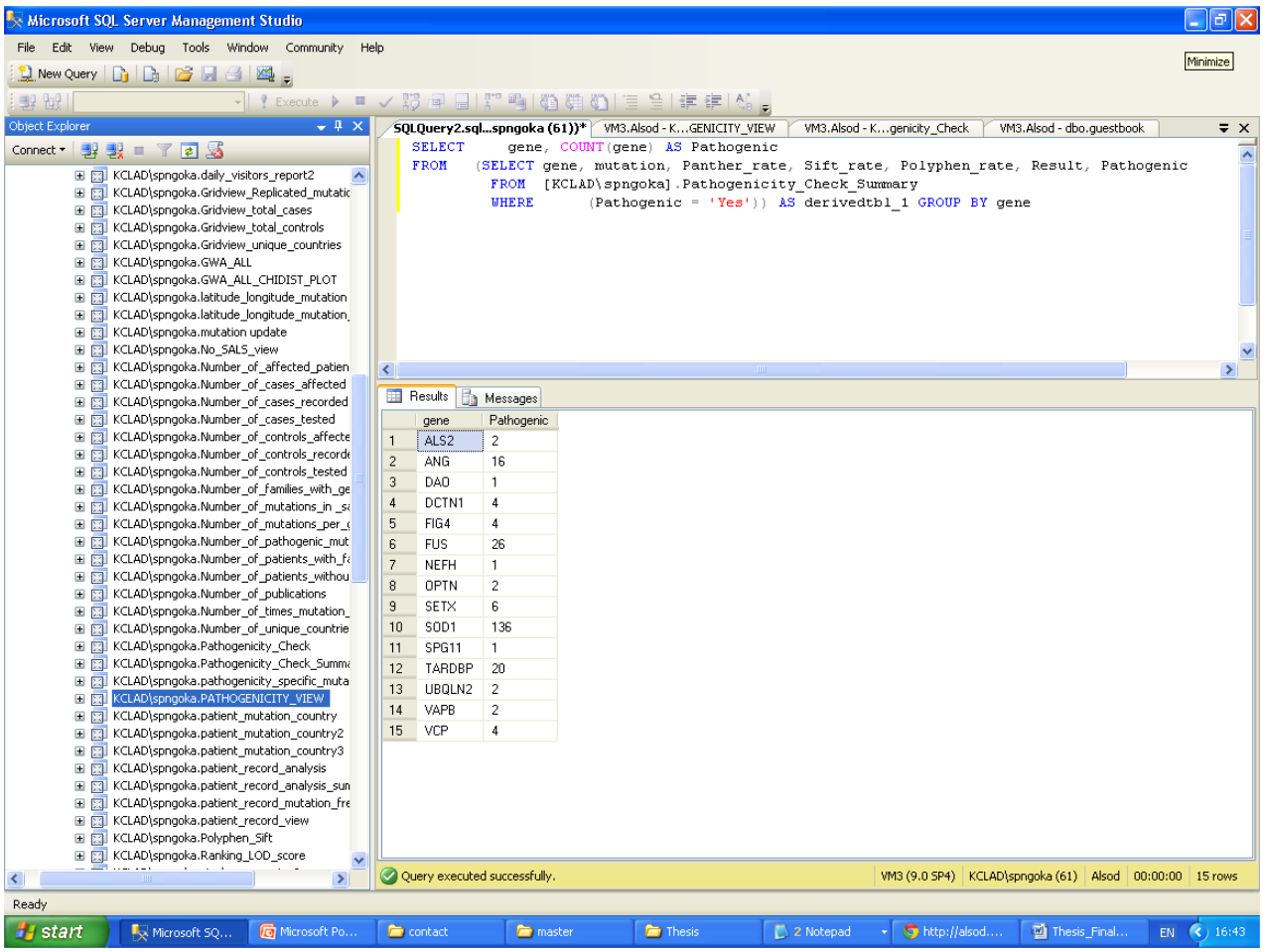


Figure 153: Query result of predicted pathogenic FALS genes mutations

For the pathogenicity prediction, using a threshold score >1 (that is, where the combination score is 2 or 3) to define pathogenicity, just 110 mutations out of 531 were identified as pathogenic, with particularly poor predictions for *FUS* and *TARDBP* when compared with biological evidence of pathogenicity. Using a threshold score of > 0 (that is, where the combination score is 1 or 2 or 3) to define pathogenicity brought the number of pathogenic mutations to 227, suggesting that about 50% of recorded FALS mutations are pathogenic based on bioinformatics predictions as at March 2014.

7.3.3 Data extraction from publications

After conducting a systematic review of all publications related to ALS genetics, there were 14 genes that fulfilled the inclusion criteria for generation of credibility score at the time of the survey, and had sufficient data manually curated from publications as explained in the data extraction process above. These were *ALS2*, *FUS*, *DAO*, *VCP*, *VAPB*, *ANG*, *DCTN1*, *FIG4*, *SETX*, *SOD1*, *TARDBP*, *SPG11*, *NEFH*, and *OPTN*.

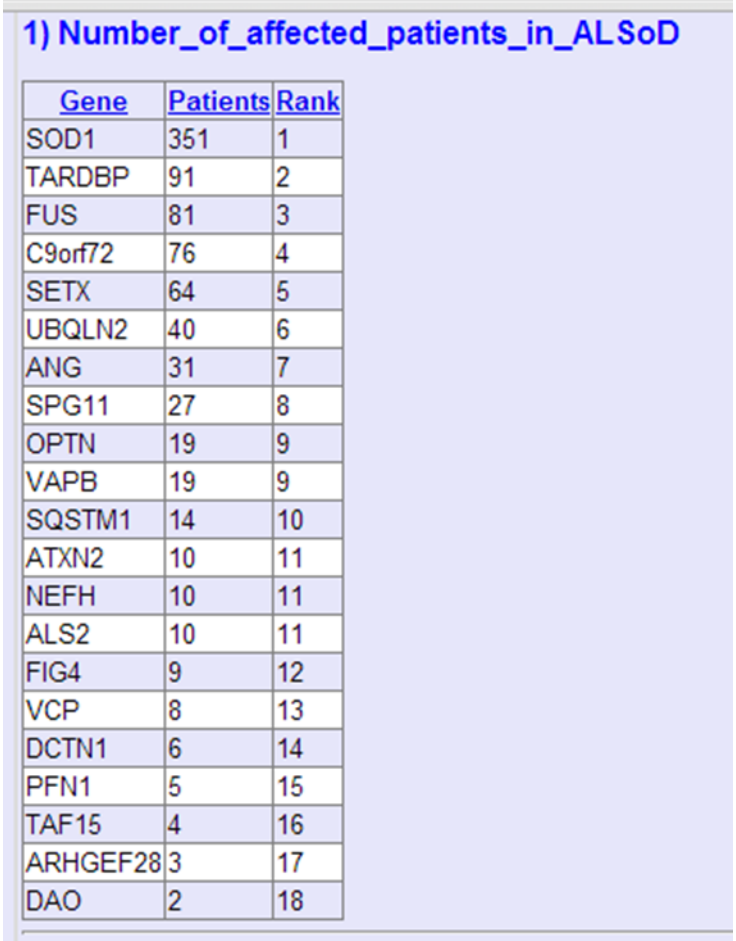
7.3.4 Automated gene ranking

Using the full set of 11 procedures, the automated method ranked these as ALS-causing genes in the following order: *SOD1*, *TARDBP*, *FUS*, *ANG*, *SPG11*, *NEFH*, *OPTN*, *ALS2*, *SETX*, *FIG4*, *VAPB*, *DCTN1*, *TAF15*, *VCP*, *DAO*.

Subsets of the 11 procedures may be defined by the user if needed. This allows flexibility in which evidence is regarded as useful. For example in Figure 3, using the number of mutations reported in a single gene and the number predicted as pathogenic as test criteria ranks the genes in the following order: *SOD1*, *TARDBP*, *FUS*, *ANG*, *OPTN*, *SETX*, *ALS2*, *SPG11*, *FIG4*, *DCTN1*, *VAPB*, *VCP*, *DAO*. The output shows that the first six genes, *SOD1*, *TARDBP*, *FUS*, *ANG*, *OPTN* and *SETX*, have a total of 121, 17, 19, 12, 5 and 4 pathogenic mutations respectively and, for example, the I113T, D90A and A4V pathogenic mutations of the *SOD1* gene were replicated in 17, 14 and 12 studies. It also shows there are 6 different mutations in codon 93 of *SOD1* and 5 different mutations in codon 521 of *FUS*. Other displayed information includes the number of countries in which gene mutations have been reported. For example, *SOD1* mutation has been reported in 34 countries with representation from every continent of the world, while *TARDBP*, *ALS2*, *ANG*, *FUS*, *SETX* and *NEFH* have been reported in 13, 9, 7, 7, 6 and 5 unique countries respectively. Genes like *FIG4*, *DPP6*, *DCTN1*, *UBQLN2*, *TAF15* which were recorded in only 1 country each have the lowest ranks.

7.3.4.1 Query generation in SQL

7.3.4.1.1 Number_of_affected_patients_in_ALSoD



1) Number_of_affected_patients_in_ALSoD

Gene	Patients	Rank
SOD1	351	1
TARDBP	91	2
FUS	81	3
C9orf72	76	4
SETX	64	5
UBQLN2	40	6
ANG	31	7
SPG11	27	8
OPTN	19	9
VAPB	19	9
SQSTM1	14	10
ATXN2	10	11
NEFH	10	11
ALS2	10	11
FIG4	9	12
VCP	8	13
DCTN1	6	14
PFN1	5	15
TAF15	4	16
ARHGEF28	3	17
DAO	2	18

Figure 154: Number of affected patients

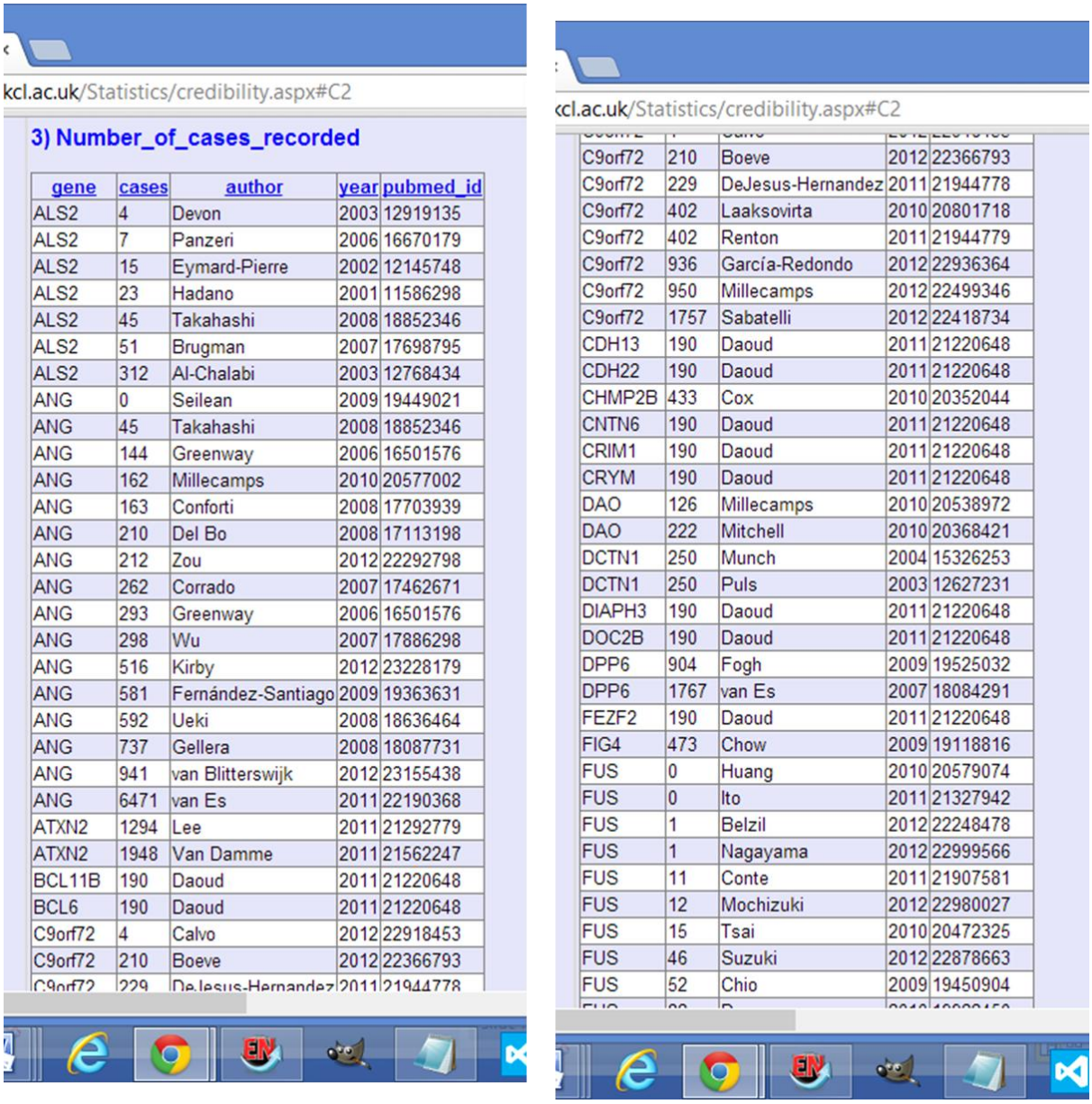
7.3.4.1.2 Number_of_mutations_per_gene

2) Number_of_mutations_per_gene

Gene	Mutations	Rank
SOD1	173	1
FUS	65	2
TARDBP	50	3
OPTN	31	4
ANG	29	5
ALS2	23	6
SQSTM1	14	7
SPG11	12	8
NEFH	11	9
FIG4	10	10
TAF15	7	11
SETX	7	11
DCTN1	6	12
DIAPH3	6	12
CDH13	5	13
UBQLN2	5	13
VCP	4	14
PFN1	4	14
OMA1	3	15
SYT9	3	15
CDH22	3	15
CHMP2B	2	16
CNTN6	2	16
ARHGEF28	2	16
BCL11B	2	16
DAO	2	16
RAMP3	2	16
VAPB	2	16
RNASE2	1	17
PCP4	1	17
NETO1	1	17
SIGMAR1	1	17
SOX5	1	17
GRB14	1	17
LUM	1	17
DOC2B	1	17
EWSR1	1	17
FEZF2	1	17
BCL6	1	17
C9orf72	1	17
ATXN2	1	17
CRIM1	1	17
CRYM	1	17

Figure 155: Number of mutations per gene

7.3.4.1.3 Number_of_cases_recorded



kcl.ac.uk/Statistics/credibility.aspx#C2

3) Number_of_cases_recorded

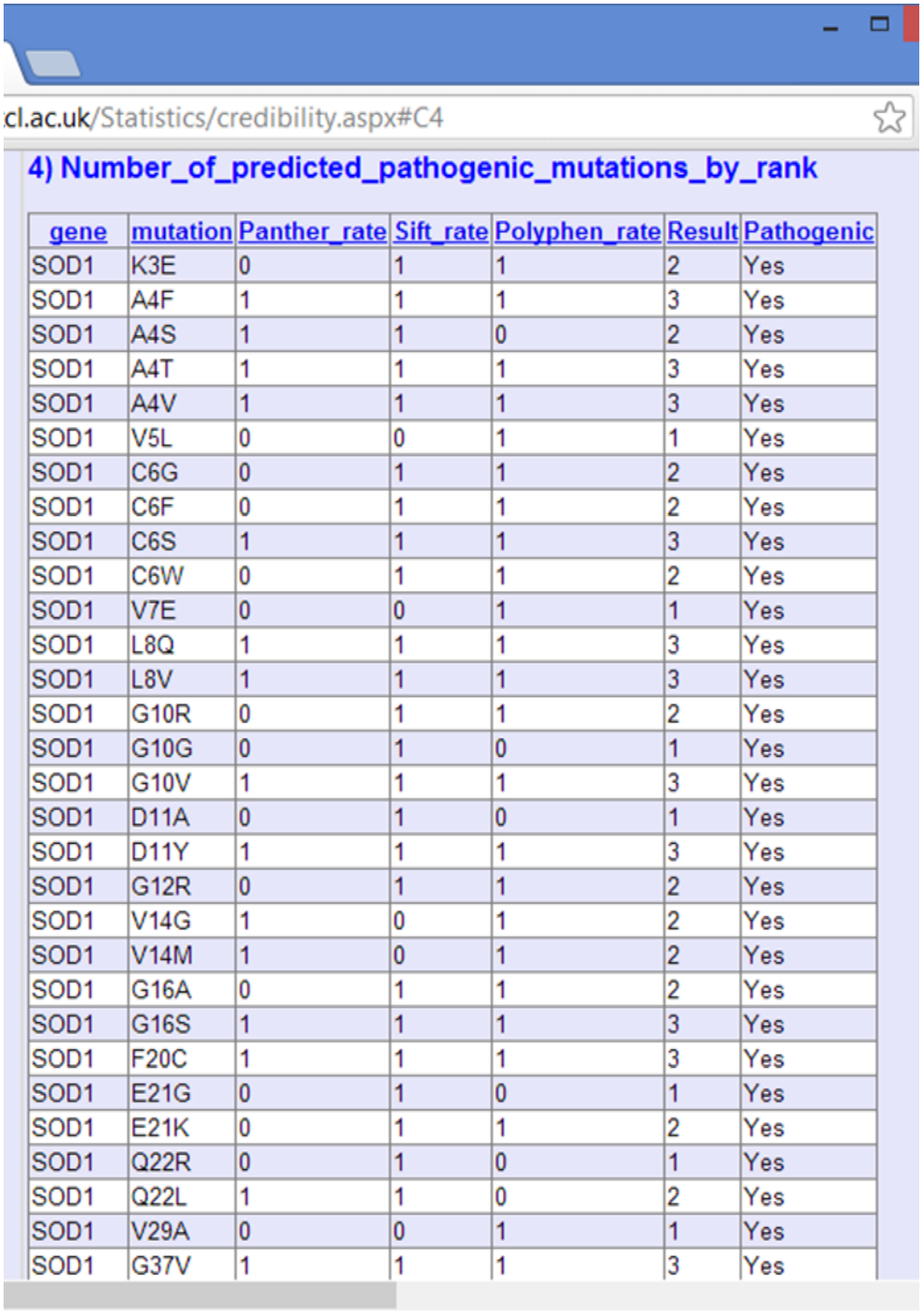
gene	cases	author	year	pubmed_id
ALS2	4	Devon	2003	12919135
ALS2	7	Panzeri	2006	16670179
ALS2	15	Eymard-Pierre	2002	12145748
ALS2	23	Hadano	2001	11586298
ALS2	45	Takahashi	2008	18852346
ALS2	51	Brugman	2007	17698795
ALS2	312	Al-Chalabi	2003	12768434
ANG	0	Seilean	2009	19449021
ANG	45	Takahashi	2008	18852346
ANG	144	Greenway	2006	16501576
ANG	162	Millecamps	2010	20577002
ANG	163	Conforti	2008	17703939
ANG	210	Del Bo	2008	17113198
ANG	212	Zou	2012	22292798
ANG	262	Corrado	2007	17462671
ANG	293	Greenway	2006	16501576
ANG	298	Wu	2007	17886298
ANG	516	Kirby	2012	23228179
ANG	581	Fernández-Santiago	2009	19363631
ANG	592	Ueki	2008	18636464
ANG	737	Gellera	2008	18087731
ANG	941	van Blitterswijk	2012	23155438
ANG	6471	van Es	2011	22190368
ATXN2	1294	Lee	2011	21292779
ATXN2	1948	Van Damme	2011	21562247
BCL11B	190	Daoud	2011	21220648
BCL6	190	Daoud	2011	21220648
C9orf72	4	Calvo	2012	22918453
C9orf72	210	Boeve	2012	22366793
C9orf72	229	DeJesus-Hernandez	2011	21944778

kcl.ac.uk/Statistics/credibility.aspx#C2

C9orf72	210	Boeve	2012	22366793
C9orf72	229	DeJesus-Hernandez	2011	21944778
C9orf72	402	Laaksovirta	2010	20801718
C9orf72	402	Renton	2011	21944779
C9orf72	936	García-Redondo	2012	22936364
C9orf72	950	Millecamps	2012	22499346
C9orf72	1757	Sabatelli	2012	22418734
CDH13	190	Daoud	2011	21220648
CDH22	190	Daoud	2011	21220648
CHMP2B	433	Cox	2010	20352044
CNTN6	190	Daoud	2011	21220648
CRIM1	190	Daoud	2011	21220648
CRYM	190	Daoud	2011	21220648
DAO	126	Millecamps	2010	20538972
DAO	222	Mitchell	2010	20368421
DCTN1	250	Munch	2004	15326253
DCTN1	250	Puls	2003	12627231
DIAPH3	190	Daoud	2011	21220648
DOC2B	190	Daoud	2011	21220648
DPP6	904	Fogh	2009	19525032
DPP6	1767	van Es	2007	18084291
FEZF2	190	Daoud	2011	21220648
FIG4	473	Chow	2009	19118816
FUS	0	Huang	2010	20579074
FUS	0	Ito	2011	21327942
FUS	1	Belzil	2012	22248478
FUS	1	Nagayama	2012	22999566
FUS	11	Conte	2011	21907581
FUS	12	Mochizuki	2012	22980027
FUS	15	Tsai	2010	20472325
FUS	46	Suzuki	2012	22878663
FUS	52	Chio	2009	19450904

Figure 156: Number of cases recorded

7.3.4.1.4 Number_of_predicted_pathogenic_mutations_by_rank



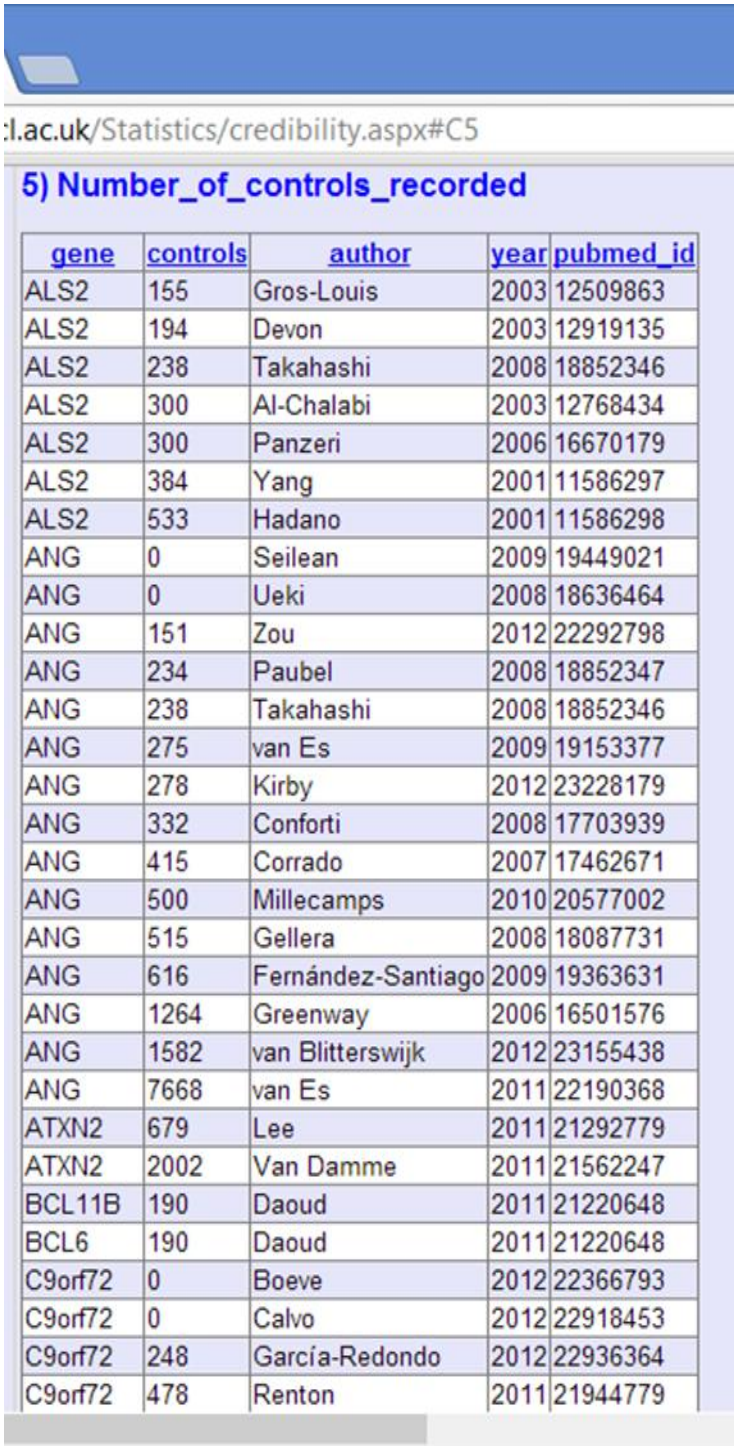
cl.ac.uk/Statistics/credibility.aspx#C4

4) Number_of_predicted_pathogenic_mutations_by_rank

gene	mutation	Panther_rate	Sift_rate	Polyphen_rate	Result	Pathogenic
SOD1	K3E	0	1	1	2	Yes
SOD1	A4F	1	1	1	3	Yes
SOD1	A4S	1	1	0	2	Yes
SOD1	A4T	1	1	1	3	Yes
SOD1	A4V	1	1	1	3	Yes
SOD1	V5L	0	0	1	1	Yes
SOD1	C6G	0	1	1	2	Yes
SOD1	C6F	0	1	1	2	Yes
SOD1	C6S	1	1	1	3	Yes
SOD1	C6W	0	1	1	2	Yes
SOD1	V7E	0	0	1	1	Yes
SOD1	L8Q	1	1	1	3	Yes
SOD1	L8V	1	1	1	3	Yes
SOD1	G10R	0	1	1	2	Yes
SOD1	G10G	0	1	0	1	Yes
SOD1	G10V	1	1	1	3	Yes
SOD1	D11A	0	1	0	1	Yes
SOD1	D11Y	1	1	1	3	Yes
SOD1	G12R	0	1	1	2	Yes
SOD1	V14G	1	0	1	2	Yes
SOD1	V14M	1	0	1	2	Yes
SOD1	G16A	0	1	1	2	Yes
SOD1	G16S	1	1	1	3	Yes
SOD1	F20C	1	1	1	3	Yes
SOD1	E21G	0	1	0	1	Yes
SOD1	E21K	0	1	1	2	Yes
SOD1	Q22R	0	1	0	1	Yes
SOD1	Q22L	1	1	0	2	Yes
SOD1	V29A	0	0	1	1	Yes
SOD1	G37V	1	1	1	3	Yes

Figure 157: Number of predicted pathogenic mutations by rank

7.3.4.1.5 Number_of_controls_recorded



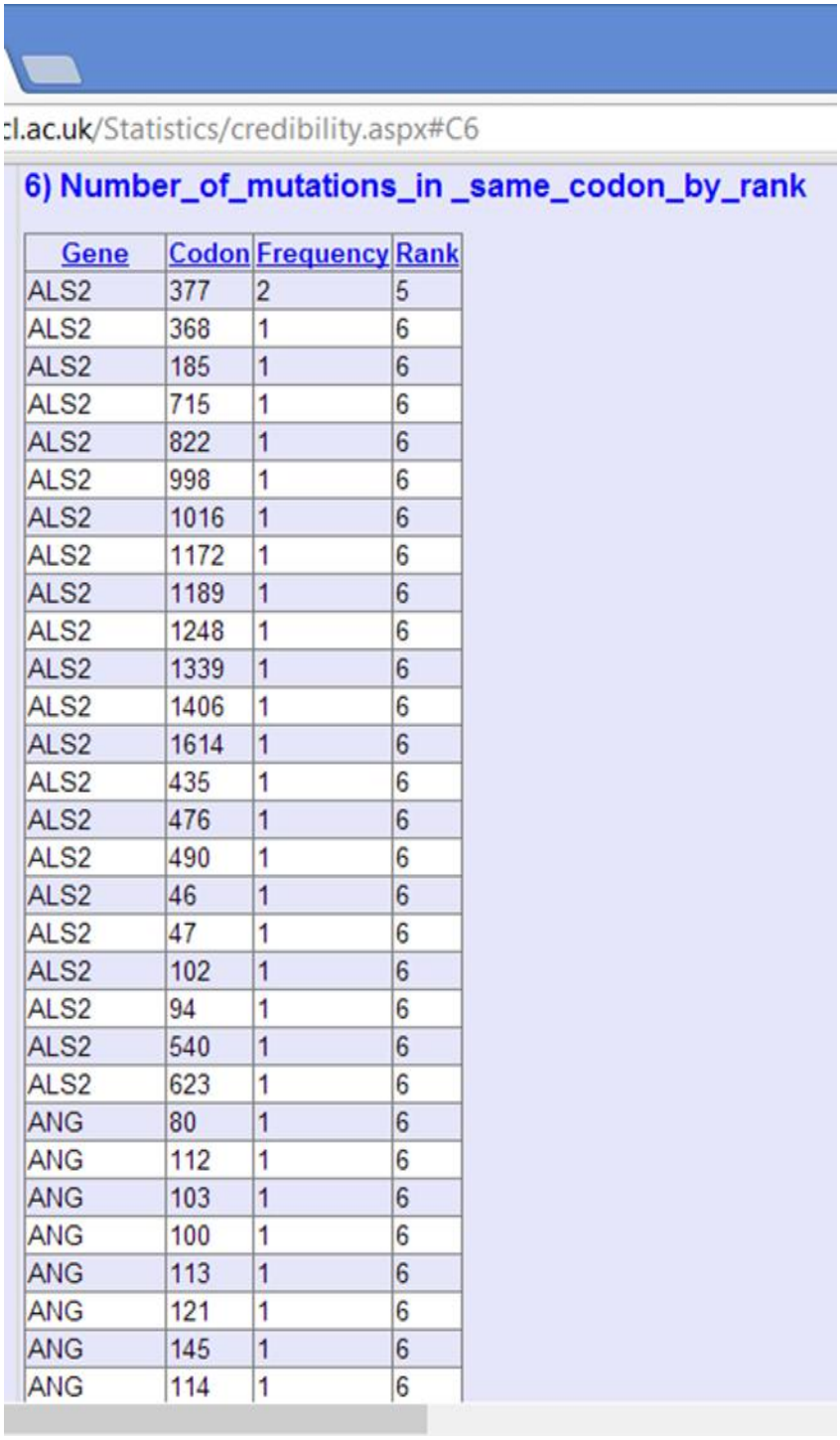
:l.ac.uk/Statistics/credibility.aspx#C5

5) Number_of_controls_recorded

gene	controls	author	year	pubmed_id
ALS2	155	Gros-Louis	2003	12509863
ALS2	194	Devon	2003	12919135
ALS2	238	Takahashi	2008	18852346
ALS2	300	Al-Chalabi	2003	12768434
ALS2	300	Panzeri	2006	16670179
ALS2	384	Yang	2001	11586297
ALS2	533	Hadano	2001	11586298
ANG	0	Seilean	2009	19449021
ANG	0	Ueki	2008	18636464
ANG	151	Zou	2012	22292798
ANG	234	Paubel	2008	18852347
ANG	238	Takahashi	2008	18852346
ANG	275	van Es	2009	19153377
ANG	278	Kirby	2012	23228179
ANG	332	Conforti	2008	17703939
ANG	415	Corrado	2007	17462671
ANG	500	Millecamps	2010	20577002
ANG	515	Gellera	2008	18087731
ANG	616	Fernández-Santiago	2009	19363631
ANG	1264	Greenway	2006	16501576
ANG	1582	van Blitterswijk	2012	23155438
ANG	7668	van Es	2011	22190368
ATXN2	679	Lee	2011	21292779
ATXN2	2002	Van Damme	2011	21562247
BCL11B	190	Daoud	2011	21220648
BCL6	190	Daoud	2011	21220648
C9orf72	0	Boeve	2012	22366793
C9orf72	0	Calvo	2012	22918453
C9orf72	248	García-Redondo	2012	22936364
C9orf72	478	Renton	2011	21944779

Figure 158: Number of controls recorded

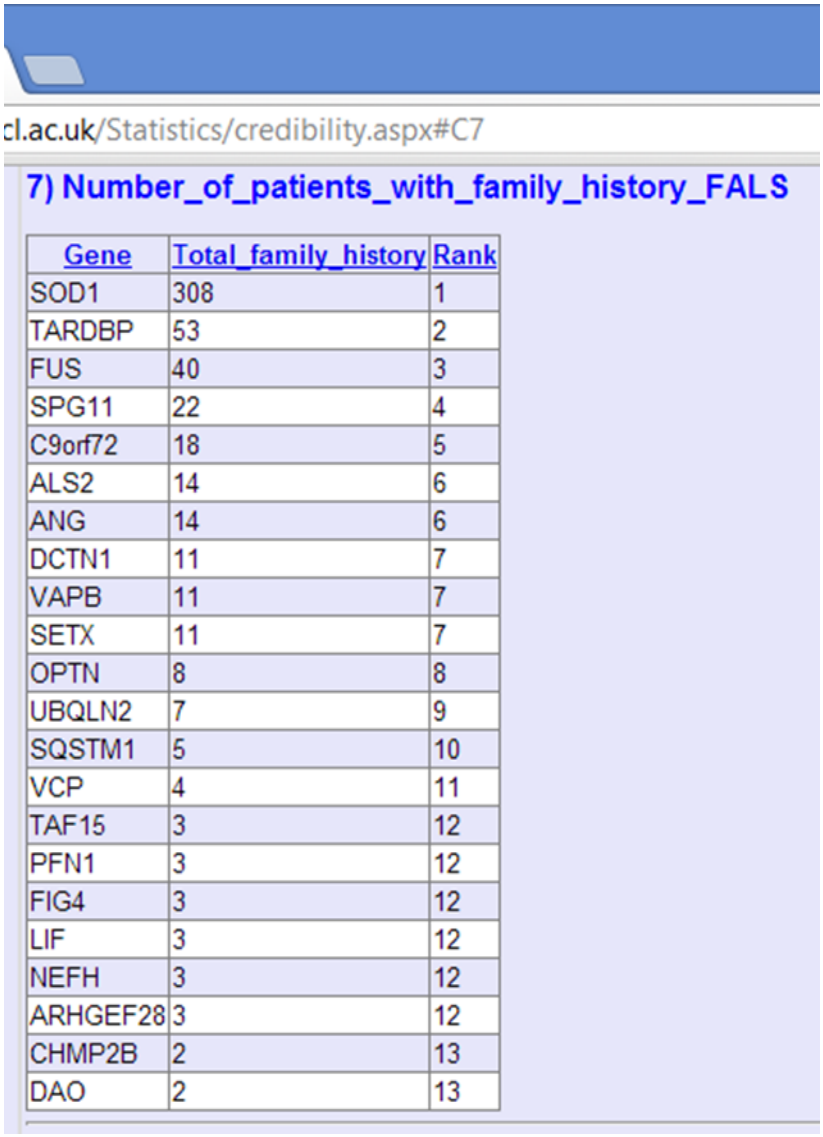
7.3.4.1.6 Number_of_mutations_in _same_codon_by_rank



Gene	Codon	Frequency	Rank
ALS2	377	2	5
ALS2	368	1	6
ALS2	185	1	6
ALS2	715	1	6
ALS2	822	1	6
ALS2	998	1	6
ALS2	1016	1	6
ALS2	1172	1	6
ALS2	1189	1	6
ALS2	1248	1	6
ALS2	1339	1	6
ALS2	1406	1	6
ALS2	1614	1	6
ALS2	435	1	6
ALS2	476	1	6
ALS2	490	1	6
ALS2	46	1	6
ALS2	47	1	6
ALS2	102	1	6
ALS2	94	1	6
ALS2	540	1	6
ALS2	623	1	6
ANG	80	1	6
ANG	112	1	6
ANG	103	1	6
ANG	100	1	6
ANG	113	1	6
ANG	121	1	6
ANG	145	1	6
ANG	114	1	6

Figure 159: Number of mutations in the same codon by rank

7.3.4.1.7 Number_of_patients_with_family_history_FALS



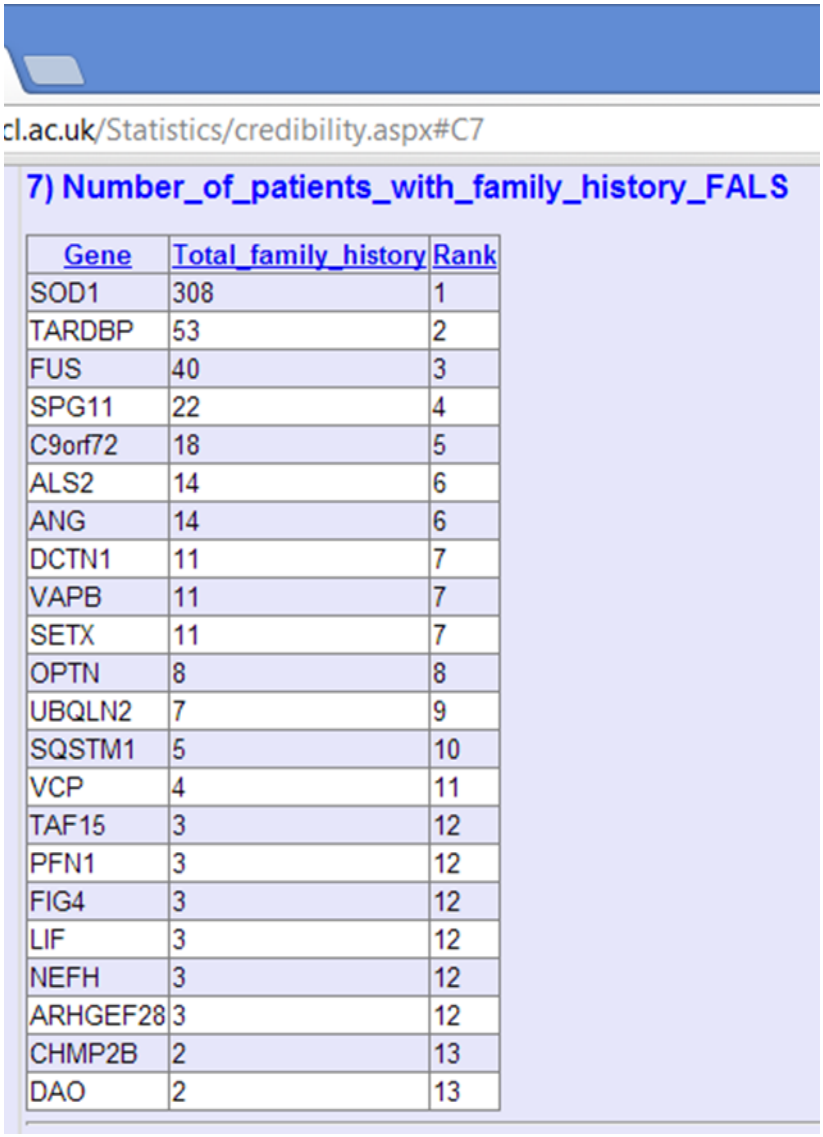
cl.ac.uk/Statistics/credibility.aspx#C7

7) Number_of_patients_with_family_history_FALS

Gene	Total_family_history	Rank
SOD1	308	1
TARDBP	53	2
FUS	40	3
SPG11	22	4
C9orf72	18	5
ALS2	14	6
ANG	14	6
DCTN1	11	7
VAPB	11	7
SETX	11	7
OPTN	8	8
UBQLN2	7	9
SQSTM1	5	10
VCP	4	11
TAF15	3	12
PFN1	3	12
FIG4	3	12
LIF	3	12
NEFH	3	12
ARHGEF28	3	12
CHMP2B	2	13
DAO	2	13

Figure 160: Number of patients with family history

7.3.4.1.8 Number_of_patients_with_no_family_history_SALS



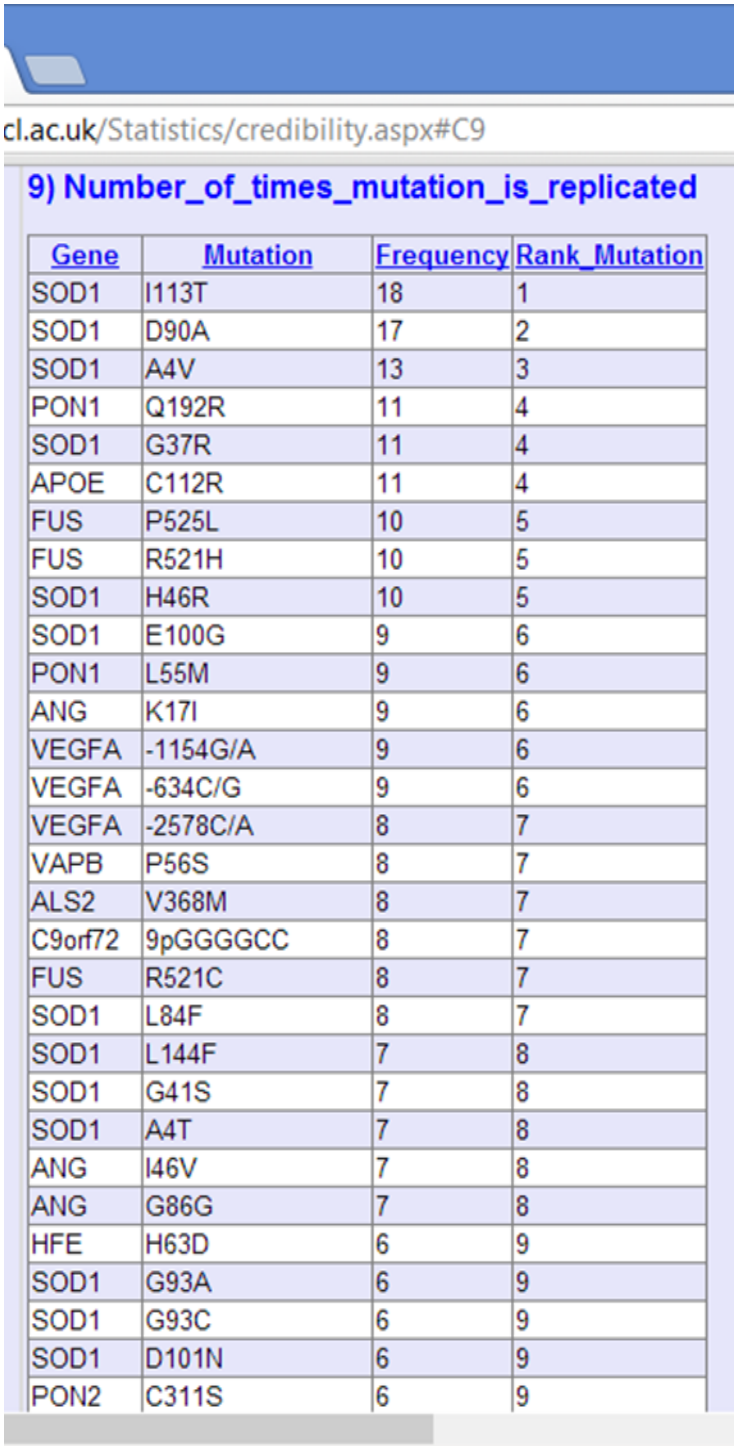
cl.ac.uk/Statistics/credibility.aspx#C7

7) Number_of_patients_with_family_history_FALS

Gene	Total_family_history	Rank
SOD1	308	1
TARDBP	53	2
FUS	40	3
SPG11	22	4
C9orf72	18	5
ALS2	14	6
ANG	14	6
DCTN1	11	7
VAPB	11	7
SETX	11	7
OPTN	8	8
UBQLN2	7	9
SQSTM1	5	10
VCP	4	11
TAF15	3	12
PFN1	3	12
FIG4	3	12
LIF	3	12
NEFH	3	12
ARHGEF28	3	12
CHMP2B	2	13
DAO	2	13

Figure 161: Number of patients with no family history

7.3.4.1.9 Number_of_times_mutation_is_replicated



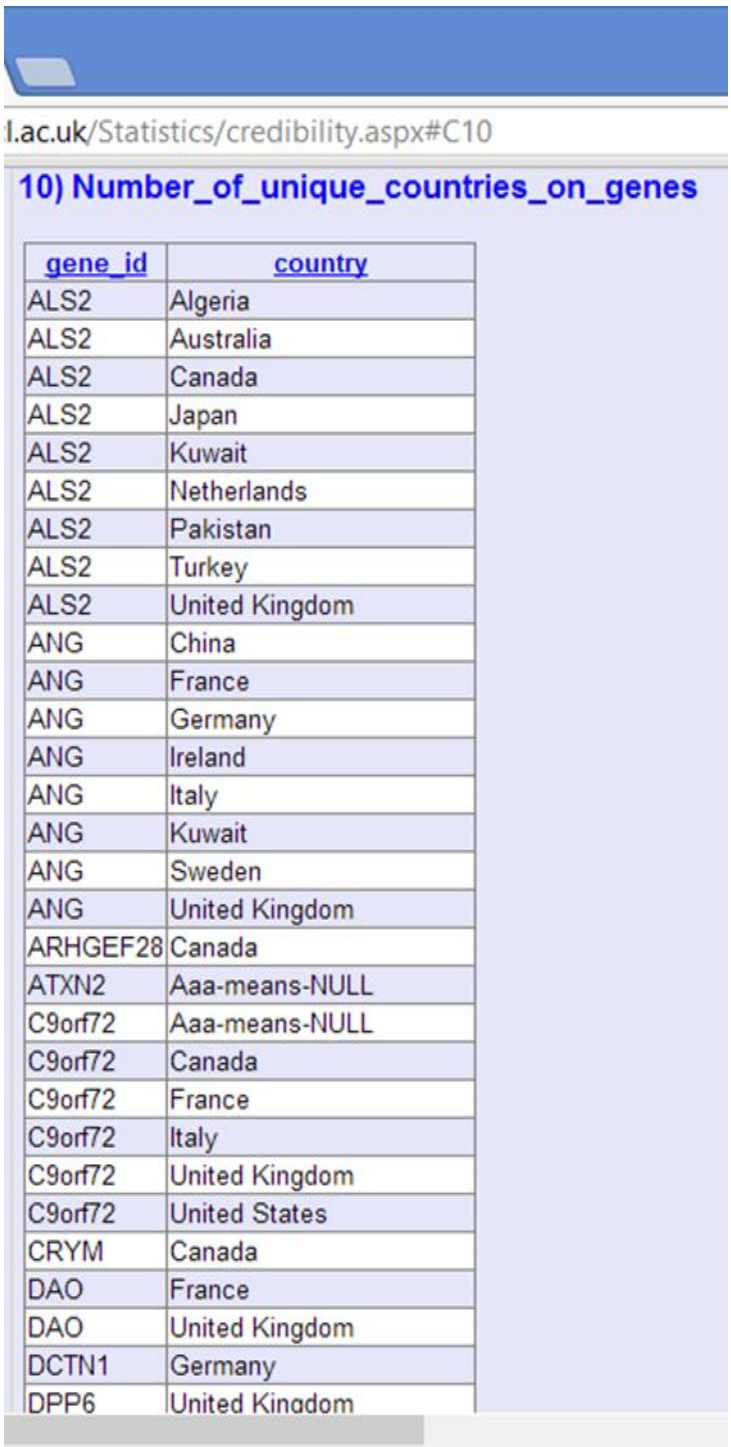
cl.ac.uk/Statistics/credibility.aspx#C9

9) Number_of_times_mutation_is_replicated

Gene	Mutation	Frequency	Rank_Mutation
SOD1	I113T	18	1
SOD1	D90A	17	2
SOD1	A4V	13	3
PON1	Q192R	11	4
SOD1	G37R	11	4
APOE	C112R	11	4
FUS	P525L	10	5
FUS	R521H	10	5
SOD1	H46R	10	5
SOD1	E100G	9	6
PON1	L55M	9	6
ANG	K17I	9	6
VEGFA	-1154G/A	9	6
VEGFA	-634C/G	9	6
VEGFA	-2578C/A	8	7
VAPB	P56S	8	7
ALS2	V368M	8	7
C9orf72	9pGGGGCC	8	7
FUS	R521C	8	7
SOD1	L84F	8	7
SOD1	L144F	7	8
SOD1	G41S	7	8
SOD1	A4T	7	8
ANG	I46V	7	8
ANG	G86G	7	8
HFE	H63D	6	9
SOD1	G93A	6	9
SOD1	G93C	6	9
SOD1	D101N	6	9
PON2	C311S	6	9

Figure 162: Number of times a mutation is replicated

7.3.4.1.10 Number_of_unique_countries_on_genes



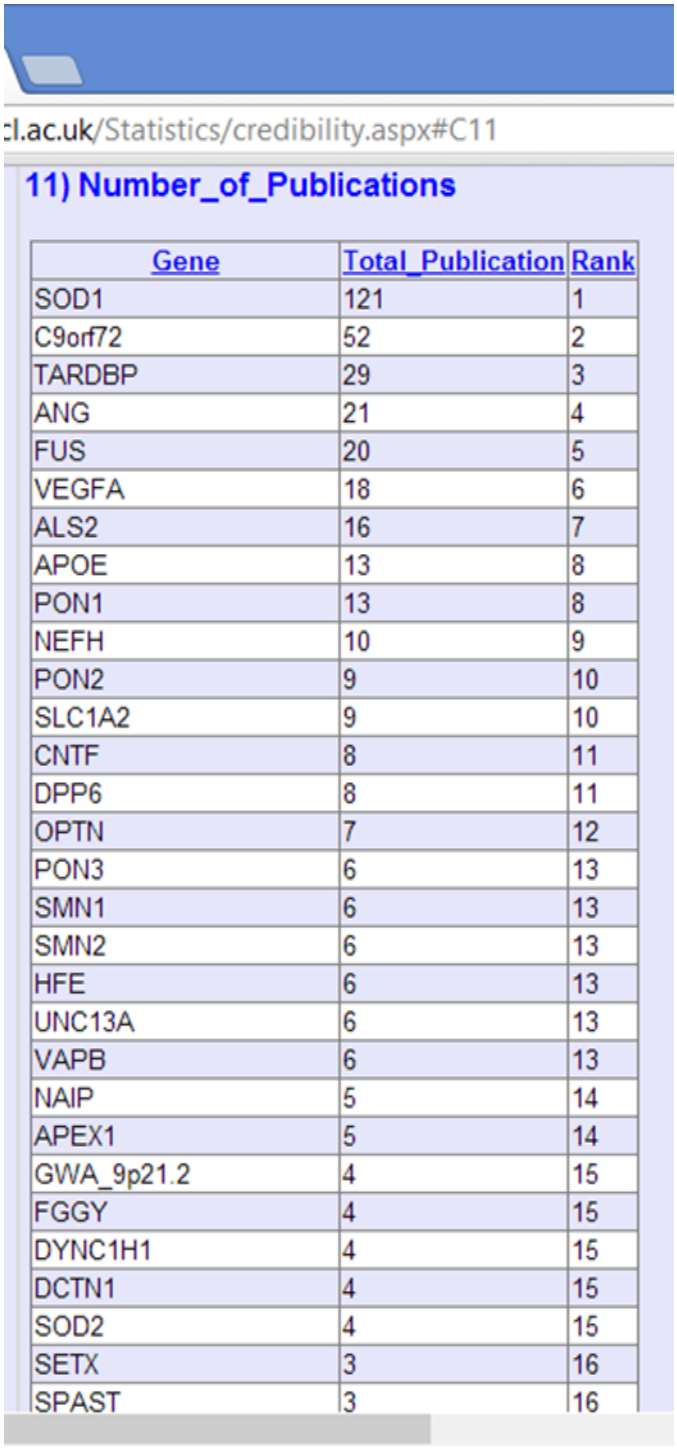
l.ac.uk/Statistics/credibility.aspx#C10

10) Number_of_unique_countries_on_genes

gene_id	country
ALS2	Algeria
ALS2	Australia
ALS2	Canada
ALS2	Japan
ALS2	Kuwait
ALS2	Netherlands
ALS2	Pakistan
ALS2	Turkey
ALS2	United Kingdom
ANG	China
ANG	France
ANG	Germany
ANG	Ireland
ANG	Italy
ANG	Kuwait
ANG	Sweden
ANG	United Kingdom
ARHGEF28	Canada
ATXN2	Aaa-means-NULL
C9orf72	Aaa-means-NULL
C9orf72	Canada
C9orf72	France
C9orf72	Italy
C9orf72	United Kingdom
C9orf72	United States
CRYM	Canada
DAO	France
DAO	United Kingdom
DCTN1	Germany
DPP6	United Kingdom

Figure 163: Number of unique countries on genes

7.3.4.1.11 Number of publications



11) Number_of_Publications

Gene	Total_Publication	Rank
SOD1	121	1
C9orf72	52	2
TARDBP	29	3
ANG	21	4
FUS	20	5
VEGFA	18	6
ALS2	16	7
APOE	13	8
PON1	13	8
NEFH	10	9
PON2	9	10
SLC1A2	9	10
CNTF	8	11
DPP6	8	11
OPTN	7	12
PON3	6	13
SMN1	6	13
SMN2	6	13
HFE	6	13
UNC13A	6	13
VAPB	6	13
NAIP	5	14
APEX1	5	14
GWA_9p21.2	4	15
FGGY	4	15
DYNC1H1	4	15
DCTN1	4	15
SOD2	4	15
SETX	3	16
SPAST	3	16

Figure 164: Number of publications

7.3.4.2 Ranking criteria

The rank score for each query was summed to generate an overall rank for the gene under study. For example, from Figure 3, the last row for the *DAO* gene gives the column score 15 for Rank_Mutations, 14 for Rank_Patients and 9 for Rank_Pathogenicity. This produces a total of 38 (that is 15 + 14 + 9) in the

Rank_Sum column. The generated Rank_Sum for all the genes are arranged in ascending order placing *DAO* 12th by final rank. On the other hand, *FUS* is placed 3rd by final rank as the corresponding scores are 3 + 3 + 2 = 8.

7.3.5 Validation of the method

8/25 ALS genetics experts selected based on having published at least one paper on ALS genetics responded.

7.3.6 User interface

The Credibility Analysis page at (http://alsod.iop.kcl.ac.uk/Statistics/credibility.aspx) allows criteria to be selected by users in the form of checkboxes. Clicking the 'Analyse' button then displays the ranked result. A detailed summary of ranked credibility data are also displayed for further reference by users giving the outcome of each procedure based query. Any combination of queries can be included in generating the score except Number of patients and Number of mutations found in each gene which are mandatory selections.

7.3.7 Statistical methods

A high Cronbach's Alpha of 0.8 shows that there's a high reliability and correlation between the three methods of testing credibility as seen in Figure 165. Comparison of the full automated method with the ALS genetics experts' rankings gave a Spearman's Rho of 0.69 (P = 0.009) for the forced expert rankings, and 0.57 (P = 0.042) for the unforced rankings, indicating a good correlation between the methods.

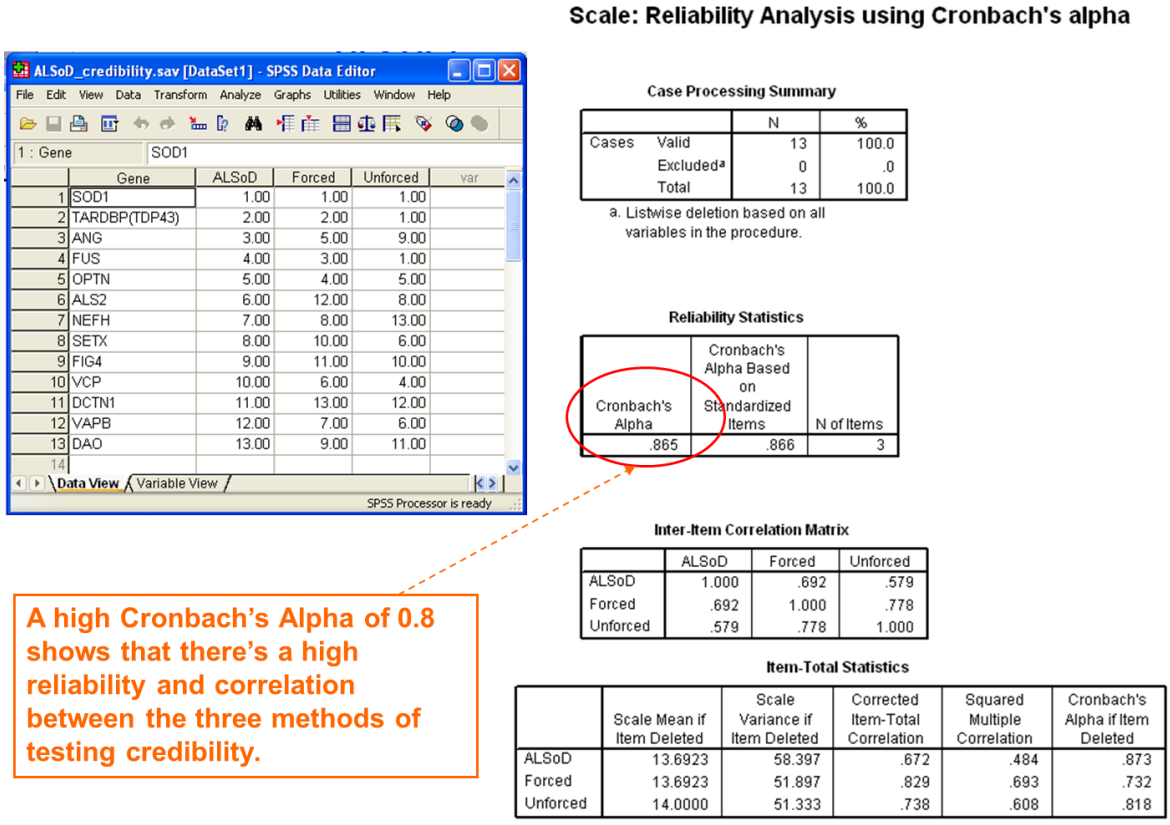


Figure 165: Reliability Analysis using Cronbach's alpha

7.4 Development of a smartphone app for a genetics website, ALSoD

7.4.1 Searched online for direction

A list of useful websites describing how to develop a mobile website ranging from third-party softwares (done by a click of a button) to developer's plug-in (developed from scratch by a programmer).

7.4.2 Choosing between a mobile website or an app

I eventually decided to develop a mobile website first. The outcome of the mobile website across multiple devices inspired me to develop an application for a platform like Android. Factors determining my decision were Target Audience (The focus on users of the mobile app is researchers and patients); Budget Available (We have a £0 budget); Purpose Intended (For easy access of the website across multiple mobile devices) and Features Required (The mobile app requires a user to have a smartphone and/or tablet and an internet connection).

7.4.3 Optimization of Webpages

In the 6-month data analysis period from August 2012 to January 2013, there were 5051 visits to the website, of which 2698 were unique (53%). There were 19,785 page views, of which 8883 (45%) were to 4

sets of pages. These pages focused on the pathogenicity of mutations, gene information, data analysis of mutations, and patient data.

7.4.4 Design Heuristics

Pages were optimized for mobile browsing by reducing image size to 5% of the original size, creating a mobile master page different from the desktop master page, and creating a link page to allow users to switch seamlessly between the mobile and desktop views.

7.4.5 Mobile Device Detection

If the UserAgent string contained keywords suggesting a mobile platform, for example, BlackBerry, Palm, mobile, iPhone, or iPad, then the user's device was redirected to the mobile site [36] displaying the compact version (Figure 125) of the website instead of the full version (Figure 44) [10].

7.4.6 Requesting Responses

To test this, I sent the mobile site URL to mobile phones of 14 users (colleagues and friends from whom I could easily obtain verbal feedback): 3 on the Android platform, 1 on the Windows Phone, 5 on BlackBerry OS, and 5 on iOS (2 iPhone and 3 iPad). All users gave positive feedback except for the Windows Phone user who could not utilize the pages with dropdown boxes.

7.4.7 App Development

Following successful implementation of the mobile website, I began app development. One straightforward method to achieve this is to automatically convert an already-built mobile website into a native app. This is done through the "WebView" object, which is an in-app Web browser used to display a website as if viewed on the browser of an Android smartphone [10]. For testing, I downloaded and used Android simulators. The plugins allow programmers to develop, test, and debug a Java application using the Eclipse IDE, but it requires a high level of programming skill [24]. We also tested and manipulated the .apk file on a real Android phone before submitting to Google Play.

7.4.8 Creating Awareness

Following the development of the mobile version, on the ALSoD Facebook page, 34 users recommended the website by May 29, 2013, using the Facebook "Recommend" button embedded on the website. Current tabular data are available on the website [37], displaying the growth of visits to the genetic database, as seen graphically in Figure 127.

7.4.9 Feedback From App Users

After the creation, testing, and publicity of the app, we received feedback from users about: caching for offline viewing [25-27], which would enable users to continue work; having a “page loading” icon when connecting; making users aware of the cookies policy; using an option menu button [28] to display analysis webpages (interaction.aspx, credibility.aspx, analysis.aspx); and creating a link to allow users to switch from mobile view to desktop view, as this would be useful on tablets like the iPad. I was able to implement all changes except for the offline viewing, which is difficult to implement because the database is large and held online.

7.4.10 Analysis of Visits

Our Google Analytics account showed that visits to the website increased from 2231 to 2820, yielding a 26% increase from pre-mobile period to post-mobile period and a 230% increase on the use of mobile devices (including tablets) to access the ALSoD website. On average, there were 300 unique visitors a day suggesting a high demand from the research community. A total of 1595 unique visitors in the post-mobile era accessed 11,376 page views on the website as opposed to 1220 unique visitors in the pre-mobile period (an increase of 31%), showing the relevance of a mobile-friendly website. Five mobile operating systems (Android, iOS, BlackBerry, Windows Phone, Symbian) were detected to have accessed the website within 6 months. Although BlackBerry OS visits declined from 34 to 14 visits (58%), iOS for iPhones and iPads increased from 40 to 105 visits (162%), and visits by Android devices increased from 29 to 213 visits (634%) (Table 6). The Google search engine was the most used to search for the website (see Table 7). The likely explanation for the great increase in the use of Android devices is the development and introduction of the Android app submitted to Google Play.

Table 6: Comparison of website visits between the pre-mobile and post-mobile development.

Operating system	Visits	Pages per visit	Avg. visit duration	% new visits	Bounce rate, %
Totals					
	26.40%	7.03%	1.48%	4.25	2.98
	2820 vs 2231	4.03 vs 3.77	00:03:55 vs 00:03:58	52.45 vs 54.77	47.45 vs 48.90

Windows							
	01/Nov/2012 31/Jan/2013	-	1945	4.25	00:04:20	49.56	44.78
	01/Aug/2012 31/Oct/2012	-	1564	3.85	00:04:13	56.46	48.34
	% Change		24.36	10.47	2.71	-12.21	-7.36
Macintosh							
	01/Nov/2012 31/Jan/2013	-	435	3.26	00:02:36	51.26	53.33
	01/Aug/2012 31/Oct/2012	-	490	3.27	00:02:50	51.02	54.08
	% Change		-11.22	-0.28	-8.71	0.48	-1.38
Android							
	01/Nov/2012 31/Jan/2013	-	213	2.35	00:01:29	76.53	63.85
	01/Aug/2012 31/Oct/2012	-	29	4.93	00:03:12	65.52	44.83
	% Change		634.48	-52.30	-53.69	16.80	42.43
iOS							
	01/Nov/2012 31/Jan/2013	-	105	4.52	00:02:31	82.86	48.57
	01/Aug/2012 31/Oct/2012	-	40	4.55	00:02:06	87.50	50.00
	% Change		162.50	-0.58	19.77	-5.31	-2.86

Linux							
	01/Nov/2012 31/Jan/2013	-	85	6.51	00:08:11	28.24	29.41
	01/Aug/2012 31/Oct/2012	-	61	3.23	00:03:13	34.43	32.79
	% Change		39.34	101.45	154.08	-17.98	-10.29
Other systems							
	01/Nov/2012 31/Jan/2013	-	15	1.07	00:00:28	100.00	93.33
	01/Aug/2012 31/Oct/2012	-	13	1.92	00:00:43	92.31	84.62
	% Change		15.38	-44.53	-35.61	8.33	10.30
BlackBerry							
	01/Nov/2012 31/Jan/2013	-	14	8.14	00:11:10	0.00	28.57
	01/Aug/2012 31/Oct/2012	-	34	7.12	00:14:12	5.88	17.65
	% Change		-58.82	14.40	-21.31	-100.00	61.90
Windows Phone							
	01/Nov/2012 31/Jan/2013	-	4	6.75	00:08:02	50.00	50.00
	01/Aug/2012 31/Oct/2012	-	0	0	00:00:00	0.00	0.00
	% Change		∞	∞	∞	∞	∞

LG						
	01/Nov/2012 - 31/Jan/2013	3	1.33	00:00:11	0.00	66.67
	01/Aug/2012 - 31/Oct/2012	0	0	00:00:00	0.00	0.00
	% Change	∞	∞	∞	0.00	∞
Samsung						
	01/Nov/2012 - 31/Jan/2013	1	1	00:00:00	100.00	100.00
	01/Aug/2012 - 31/Oct/2012	0	0	00:00:00	0.00	0.00
	% Change	∞	∞	0.00	∞	∞

Table 7: Referral traffic from search engines from August 2012 to June 2013

Source	Visits	Pages per visit	Avg. visit duration	% new visits	Bounce rate, %
Totals / Site avg					
	4506	4.09	00:04:55	42.92	46.25
	44.65%	3.68	00:04:03	51.26	51.95
Google	4339	4.06	00:04:54	42.73	46.58
Yahoo	62	3.63	00:02:42	51.61	41.94
Bing	32	5.53	00:07:11	62.50	31.25
Baidu	21	6.19	00:12:50	23.81	23.81

Daum	13	3.46	00:06:35	30.77	30.77
Ask	10	4.9	00:08:33	20.00	20.00
Conduit	8	17.62	00:08:27	37.50	25.00
AOL	7	1	00:00:00	71.43	100.00
Other search engines	5	6	00:04:24	0.00	0.00
Yandex	4	1.25	00:00:04	100.00	75.00
Babylon	3	1.33	00:00:33	100.00	66.67
AVG	1	1	00:00:00	100.00	100.00
Comcast	1	1	00:00:00	100.00	100.00

7.4.11 Analysis to date (January 2014)

From Google Scholar, the analysis of ALSoD website since August 2012 to January 2014 which is approximately 17 months has a record of 16871 visits, 8534 unique visitors and 62246 pageviews where half of the visitors are new as shown in Figure 166.

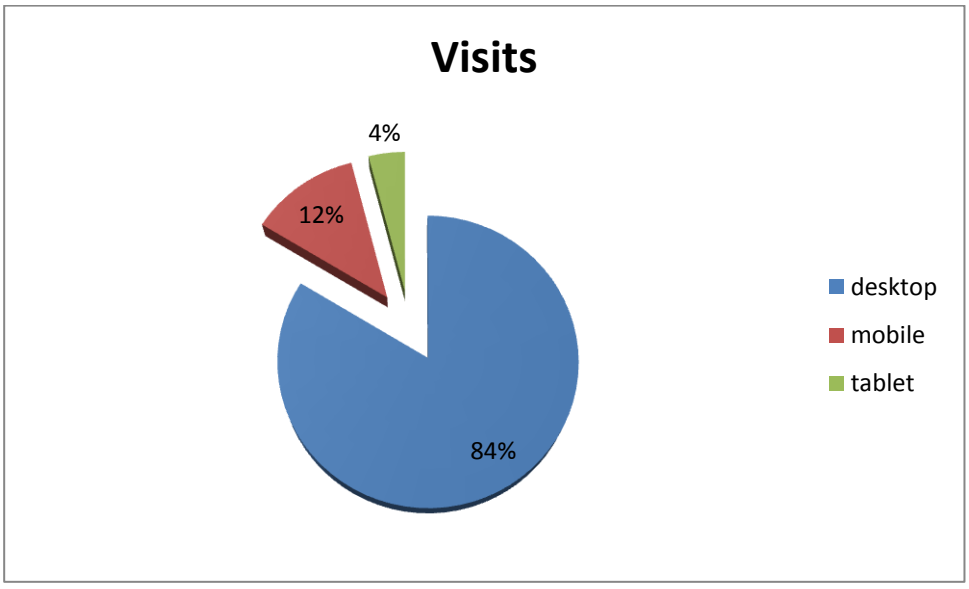


Figure 166: Visits on ALSoD for 17 months on devices

As shown in Table 8, out of 16871 visits, 456 visits were from unknown mobile devices, 329 from Apple iPad, 214 from Apple iPhone, 97 from Samsung GT-I9100 Galaxy SII, 84 from Samsung GT-I9300 Galaxy SIII, 51 from Samsung GT-N7100 Galaxy Note II and 50 from Google Nexus.

Table 8: First 10 Browser and OS used to access ALSoD from August 2012 to January 2014

		Visits	% New visits	New visits
		16,871	50.64%	8,543
1	Chrome	**5,673**	43.45%	2,465
2	Firefox	**3,402**	52.94%	1,801
3	Internet Explorer	**3,027**	63.79%	1,931
4	Safari	**2,568**	40.97%	1,052
5	Android Browser	**1,787**	61.56%	1,100
6	Opera	**193**	17.62%	34
7	Opera Mini	**38**	92.11%	35
8	Mozilla Compatible Agent	**36**	97.22%	35
9	IE with Chrome Frame	**33**	51.52%	17
10	Safari (in-app)	**31**	87.10%	27

As seen inFigure 167: First 10 Countries accessing database, the US has the largest number of visitors followed by the UK. This corresponds with the result on the ALSoD monitoring page but there is a bit of discrepancy with the third largest number of visitors from India which is quite different from the result on the in-house built monitoring page. An explanation to this could be that the IP address from India has not been recognised so it classifies visits from India as Unknown, hence the large number of Unknown countries.

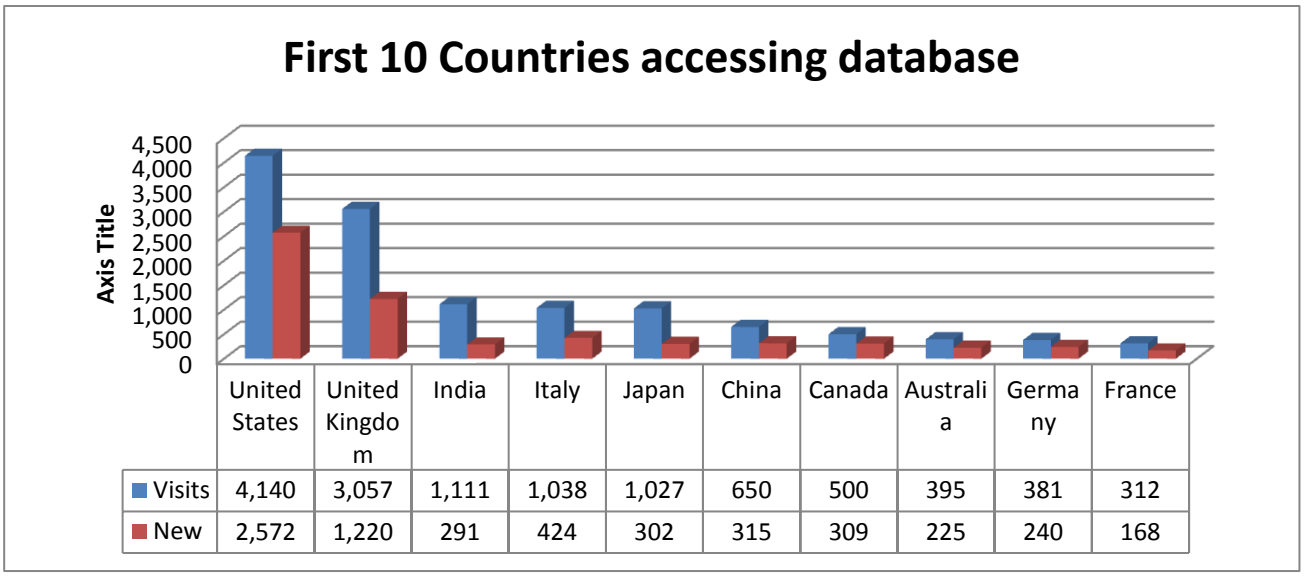


Figure 167: First 10 Countries accessing database

In Figure 168, item 1 is the main web address http://alsod.iop.kcl.ac.uk which leads a user to either the desktop homepage view on item 2 or the mobile homepage view on item 3. This result proves the high accessibility to mobile devices by users of the database.

	Page	Pageviews	% Pageviews
1	/	9,668	15.53%
2	/Index.aspx	4,141	6.65%
3	/Mobile/index.aspx	2,851	4.58%
4	/misc/dataDownload.aspx	2,789	4.48%
5	/Overview/gene.aspx?gene_id=SOD1	2,498	4.01%
6	/index.aspx	1,828	2.94%
7	/Mobile/Mutation.aspx	1,430	2.30%
8	/Statistics/pathogenicity.aspx	1,370	2.20%
9	/Overview/gene.aspx?gene_id=ANG	1,273	2.05%
10	/Statistics/analysis.aspx	1,161	1.87%

Figure 168: First 10 Pageviews

Chapter 8 DISCUSSION AND FUTURE WORK

8.1 Bioinformatics exploration of ALS genetics

Since the first discovery of mutations in the locus ALS1 of chromosome 21q22.11 in the Cu/Zn superoxide dismutase 1, soluble gene (SOD1), there has been an explosion in the number of mutations discovered in various populations. A rapid increase from 11 mutations in 1993 [91] to 43 mutations in 1996 [594] in SOD1 gene alone propelled the ALSOD Consortium to set up a uniform centralized database with the aim of pooling all available genetic and clinical data in one location [567, 594].

Due to improved technology, SOD1 mutations alone have increased from about 100 mutations in 2008 [568] to more than 170 mutations as at January 2014 on http://alsod.iop.kcl.ac.uk. Currently, over 470 mutations in more than 110 genes are recorded on the database with expectations for more genetic data to be discovered in the future. Currently, the number of disease loci associated with classical or typical ALS as at January 2014 as displayed on the website are 20 [595] while 2 loci have been reported as ALS with frontotemporal dementia (ALS-FTD) [596].

Genetic studies of ALS are often difficult for non-geneticists to understand. For geneticists, the sheer number of studies means it can be difficult to remain up-to-date. For researchers it would be helpful to have all genetic research findings in one place with tools to help understand them or to follow up the results with other analyses. The ALS Online Genetics Database, ALSoD, is a website housing the database of up-to-date ALS genetics research. For each listed gene there are links to information or analysis tools. For some genes such as SOD1 that have been extensively studied, there are additional specific links for example to predict the effect of mutations on structure and function. ALSoD was therefore developed to handle genetic data produced in ALS. Mutation and patient data were collated to understand the relationships between genotype and phenotype analysis.

Bioinformatics tools are being developed regularly to analyze these data and ALS researchers are keen on getting up-to-date information about genes, publications, mutations and analysis provided through the website. [486, 521]. ALSoD keeps up with the rapid advances in genetics by adding new gene findings as they occur and developing new tools for researchers. The remit of ALSoD was expanded to include findings from frontotemporal dementia research because there is an overlap with ALS in some cases.

The ALSoD and the ALSGene databases which were independently developed have been working together to collate both Mendelian and non-Mendelian genetically complex ALS results from up-to-date studies [486].

An independent measure of credibility that might guide non-geneticists in knowing the level of confidence to place in reported findings was developed [522]. There are several analysis tools available in ALSoD at present, including a visualiser for geographical distribution of mutations, visual tools for the display of mutations and genetic variations, and tools for reanalysis of genome-wide association data, including the ability to combine existing data with newly uploaded user data confidentially.

I developed tools to store and analyze information from genome-wide association studies, whole genome sequences and gene expression information. We implemented the largest GWAS meta-analysis in ALS based on data available in our laboratory on the ALSoD platform using the Fisher's method of combining *P*-values, which is a traditional approach [597, 598] accessible from http://alsod.iop.kcl.ac.uk/GWA2/index.aspx. There are also other analytical models like the Bayesian approach [507] which may be introduced in the near future. An On-the-fly analysis of GWAS data is carried out by combining data available on the database with unpublished user-data confidentially uploaded to carry out dynamic analysis on existing studies. The user data is formatted accordingly before they are uploaded and the result is fed back in minutes without storing users' data on the database. The most current meta-analysis on ALS patients as at December 2013 is available on the database [544]. A significant problem is how to deal with population stratification [599], but we are exploring methods using a principal components analysis approach. A more comprehensive and detailed meta-analysis of all ALS GWAS is available at www.alsgene.org. I also included GeneMANIA web interface for predicting gene interactions in ALSoD as a bioinformatics tool to forecast the functional relationship between two genes [600].

8.2 ALSoD versus Others

"ALSOD" is the original name of the database suggesting association with ALS and SOD1 gene. It is now renamed 'ALSoD'. ALSoD is an open source data repository designed for sharing clinical and genetic data. It now provides both the scientific community and wider public with up to date information on ALS-related genes. Genetic research is advancing at an extremely rapid pace. The purpose of ALSoD is to curate the available evidence for genetic risk and modifying factors for ALS with summaries of the literature and links to appropriate genetic resources. ALSoD first came online in 1999 solely for reporting SOD1 gene

mutations [567] but with the rapid advances in genetics was in need of a major overhaul and required redesigning to allow complex genetic analyses and data comparisons automatically. With the aim of providing users with maximal cross-referenced bioinformatics investigation, the redevelopment process is currently being carried out [568]. With the advent of newer technologies, international collaborations between scientists and greater awareness of research aims by those affected by ALS, what is being learnt about ALS is expanding every day.

The ALS Online Genetics Database website was initially located at http://www.alsod.org and is now located at http://alsod.iop.kcl.ac.uk. Funding is obtained from the MND Association of Great Britain and Ireland, the ALS Association, ALS Canada, MNDA Iceland and the ALS Therapy Alliance.

ALSoD allows users to submit new mutation and patient data. It also has various tables in the database schema designed for storing data on genes, mutations, patients, codons, gene sequence, trinucleotides, country details, users and stored procedures. Microsoft .NET framework with Microsoft SQL server 2008 are used to manage the database stored on the VM3 server of the Institute of Psychiatry, KCL (size about 10830 MB). Microsoft Visual Web Developer 2008 Express Edition is used to develop the web pages (current size about 1.05TB containing 387 files and 53 folders).

The second publication on ALSoD in 2008 proposed future plans for the database which is to expand and incorporate mutational data for additional genes linked to ALS; developing more graphical and statistical summary webpage and further improving the appearance and functionality of ALSoD [568]. The present PhD project has helped to achieve this plan and beyond.

Further to the discussion in 1.20.2 about other databases, ALSoD has an edge over other databases as described below.

8.2.1 Daily Updates

The usefulness of LSDBs is determined by how current the data in them is. From a study conducted in 2002, 38% of databases were last updated in the previous 6 months, 10% in the first 6 months of 2001, and 12% in 2000 [601]. A date of last update was not shown in 36% of LSDBs leaving the user confused about how current the information was [463]. Updates are carried out daily in ALSoD using scripts and internet bots. Apart from dead databases with broken links, many of the databases are updated weekly, monthly or not updated at all. I checked the Tokyo ALS mutation database (https://reseq.lifesciencedb.jp/resequence/SearchDisease.do?targetId=1) on 23rd January 2014 discovering that this website has not been updated since 18th November 2010 which is more than 3 years ago.

8.2.2 Data submission

Databases of genetic polymorphisms such as ALFRED and dbSNP rely on author submissions and contain little additional data analysis [602], whereas ALSoD has a network of contributors worldwide and a curator who administers the database regularly.

8.2.3 Missing genetic data

I compared the features of OMIM and HGMD databases and found significant inconsistencies like absent genes and missing mutations between the two databases. For example, the last check in December 2013 records 105417 mutations in the public entries (for Academics/non profit users) and 148413 mutations in the Professional domain. A search for total mutations available for the FUS gene on http://www.hgmd.cf.ac.uk/ac/gene.php?gene=FUS showed 42 mutations in the public site but 83 mutations in the professional site while on ALSoD database, there are 77 mutations available. On OMIM, 9 pathogenic FUS mutations were recorded on http://www.ncbi.nlm.nih.gov/clinvar?term=137070[MIM].

This evidence therefore strengthens the case for gene specific curation of mutations and a general plan for collection, curation, storage and publication of mutation data as in ALSoD.

8.2.4 Global representation

There are country-based mutation databases like the Iranian Human Mutation Database [603] and the Turkish Mutation Database [464, 604]. The recently developed Mutation database in Japan is also a region-centred database focusing more on data from Asia. ALSoD is unique in having data extracted from and submitted by researchers worldwide. Currently there are contributors from 24 institutions from all around the world who contribute data to ALSoD regularly.

8.2.5 Popularity of database

When accessed online on 26 June 2010, ALFRED had been visited 171,236 times since March 29 1999 while HGMD had 1,599,773 hits on its page since June 2005. That is, ALFRED is visited on an average of 15,567 times per year, and HGMD 319,955 per year. The ALSoD visitors' statistics page shows that since January 2009 (within 1.5 years), the site has been visited over 145,000 times averaging about 100,000 visits per year.

8.3 Limitations of Genetic Databases

8.3.1 Continued Funding

After much work has been carried out by curators to bring genetic databases to life, there is a great risk of loss of continued funding. Many databases are either outdated or have broken inaccessible links because the sponsorship expired and was not renewed. This could result from the lack of manpower or economic recession. ALSoD faced that situation between 2000 and 2007 which resulted in losing the domain name and website www.alsod.org.

8.3.2 Duplication of Efforts

Duplication of efforts is not impossible especially where there are hindrances to collaborating with other researchers. No curator would like to waste time 'reinventing the wheel' hence the use of hyperlinks in ALSoD to already available data on other external databases was implemented.

8.3.3 Limited resources

Resources like storage spaces will be easily used up due to high volume of genetic data produced regularly especially with advancement in technology. ALSoD keeps track of all IP Addresses visiting the website, stores all mutation and published patient data and hosts Genome Wide Association Meta Analysis data. All of these require storage space of which could pose a problem if the KCL institute hosting the database caps the storage space for each website.

8.3.4 Political boundaries on data sharing policies

Different countries have policies on data protection of its citizens. Some allow part of the data to be displayed as long as the individual remains anonymous, some require an endless list of paperwork and approval before collaborating outside their own country, some even do not allow any form of sharing at all. So, patient data in ALSoD is limited to data from publications and very few from individual researchers in countries where data-sharing restrictions are less strict.

8.4 Limitations of the ALS Online Database

8.4.1 Storage

Volume and size of data allowed for upload by the host server is limited to avoid an exhaustion of disk space provided for ALSoD on the server and to secure the accessibility of the server to external intruders trying to hack the entire system. As this tool becomes popularly used, we shall be faced with the challenge of providing a larger disk space to host this tool. A resource able to handle next generation sequencing

data and deliver useful analysis for processing and displaying such information is desirable. There is a question of whether the current ALSoD network can cope with multiple users trying to download data simultaneously. Dealing with FASTQ, BAM and VCF files to align raw data reads to the genome will be a great challenge.

8.4.2 Funding

The life line of any successful project is financial stability which involves both human and non-human resources. Many databases have come and gone while not many can boast of continuity. Broken links to websites are experienced often due to lack of maintenance. Domain names and host servers are not free and so a need for annual fees by organisations hosting websites is mandatory to keep the website alive. ALSoD went through that phase when the alsod.org domain name was lost due to lack of funds and manpower. The aspect of technological facilities has been solved by using the Institute of Psychiatry server but the sustenance of the Curator, Web Developer, Programmer, Database Administrator and/or Content Manager of the database is a pending issue.

8.4.3 Browser Incompatibility

Incompatibility issue with browsers (FireFox, GoogleChrome, Internet Explorer, Safari etc) is a technical problem beyond a developer's control. We are working on how to integrate various tools from other groups with minimal interruption from the technicalities embedded in browsers used.

8.4.4 Upgraded versions of third-party applications

An upgrade in the version of a previously used web service like PolyPhen which made a part of the results in predicting pathogenicity has become obsolete.

Also, spammers and hackers target websites regularly. I noticed in February 2014 that despite the use of Recaptcha by Google, a problem of unsolicited comments posted regularly on the feedback page resurfaced (Appendix 42). Removing the login barrier (where users provide a username and password) is the negative side to allowing the public free access to pages. Even though the software Recaptch.Net zip file was extracted, it seems to work but there is a possibility that this problem may reoccur when the current recaptcha.dll file becomes obsolete.

8.5 Concluded Research Questions

Firstly, is it possible to generate a database that summarizes genetic data for a disease and allow meta-analysis online? ALS is used as a model disease in the present study. This was done by creating gene and chromosomal overview pages of all available ALS-related genes; collating published GWAS data from the largest GWAS study of ALS to date completed by our group; collating published linkage studies and merge results on GWAS; and allowing user-defined queries of meta analysis of association studies.

Secondly, can all genetic data useful to ALS be automatically collected? Data mining from large external databases like UNIPROT (which is a consortium between the European Bioinformatics Institute (EBI), the Swiss Institute of Bioinformatics (SIB) and the Protein Information Resource (PIR)) were interrogated. Also manual curation of genetic data from publications has been carried out.

Thirdly, how achievable is an on-the-fly analysis of genetic data online? A tool for dynamically uploading GWAS data from users was generated; user data were compared with available GWAS data in the database; and user-configurable queries using a simple interface was created.

Fourthly, to what extent can the database (the ALS Online Genetics Database) be integrated with other bioinformatics resources available online? This was done by providing users of the database with maximal cross-referenced bioinformatics and investigation utilized using information from other well-known scientific and non-scientific databases like Entrez Gene, UCSC, OMIM, Genecards, KEGG, Uniprot, iHop, dbSNP, Pubmed, Wolfram Alpha, PDB, Gene Onology and others.

Lastly, with the advent of various ALS-related genes, can a database generate levels of evidence to support or refute genetic associations and linkages with ALS? Various criteria were intended to generated scores for prioritizing the importance of each gene. The initially proposed criteria for each gene are a score for the number of mutations related to ALS on ALSoD; a score for the number of published articles related to ALS available in Pubmed; a score for the number of significant SNPs within a threshold derived from the meta-analysis tool; a score for how each gene relates to other genes using Ingenuity pathway tool; and a score for the number of countries where mutations or disease-associated variants are found. This was eventually done by ranking queries on the number of affected patients, number of mutatios per gene, number of cases recorded, number of predicted pathogenin mutations, number of FALS, number of SALS, number of replicated studies and number of unique countries found.

8.6 Extra Development

In the process of developing the database, various needs arose from current developments in technological advancement, feedbacks received from users of the database and suggestions from conferences and symposiums attended. These were not included in the original research questions.

8.6.1 Development of mobile computing applications

Current trends in computing mean that users expect services to be available from any device and platform. Researchers often travel and might have a new idea while they are away from the office. It would therefore be helpful for the ALSoD data and analysis tools to be available through an application interface specifically designed for a small screen. I hence developed an application (app) for Android mobile computing systems to allow ALSoD functionality easily accessible from a smartphone or similar interface. Other developments will be carried out on iPhone and Blackberry.

8.6.2 Expansion of phenotypic inclusion criteria to include frontotemporal dementia

An interesting genetic region is on chromosome 9p21.2, identified in linkage studies of familial ALS with frontotemporal dementia (ALS-FTD), familial and sporadic ALS, and sporadic FTD [605-609]. Since there is clinical and pathological overlap of ALS and FTD it is likely that variants of the same gene are responsible for all these findings. Only three annotated genes are in the region: *c9orf72*, *IFNK* and *MOBKL2B* and there was a very active search for the disease gene [398, 610] until a study showed that a large hexanucleotide (GGGGCC) repeat expansion of C9ORF72 is present in the first intron of majority of ALS-FTD cases [405]. There is very strong evidence for the clinical and pathological overlap of ALS and FTD, hence the exclusion of FTD data from ALSoD risks leaving out important information. We therefore widened our inclusion criteria for genes, papers, and stored data to include FTD and ALS-FTD syndromes. We began the process of advertising ALSoD to the FTD community and in this way helps to establish new links to different types of information.

8.6.3 Suggestions from feedback

Users of the database have played a significant role in the development of ALSoD through valuable suggestions in the past. We depend on feedbacks through Social Media, feedback page, blogs, poster presentations, seminars, conferences, surveys etc. for further improvement.

8.7 Future of ALSoD

ALS Online Genetics Database has been considerably developed over the last decade, taking it from a broken web server with no backup, storing out of date information restricted to SOD1-induced familial ALS, and transforming it into a sleek computing engine with automated backup, optimal security, and very up-to-date information on all familial and sporadic ALS genetics, currently including several GWAS data and more than 110 genes.

The last publication on ALSoD in 2008 (before I took over the project) proposed future plans for the database which is to expand and incorporate mutational data for additional genes linked to ALS; developing more graphical and statistical summary webpage and further improving the appearance and functionality of ALSoD [6]. The database has since then exploded with up-to-date gene list which are publicly available useful tools to aid researchers in the field of ALS.

For the future of the Amyotrophic Lateral Sclerosis Online genetics Database – all things being equal – I hope to cover the subtopics explained below:

8.7.1 Collaborations

High-throughput and Next generation sequencing has transformed genetics. The quantity of data makes it difficult to handle, display and analyse with current technology. There are methods for simplifying this and as the costs drop further, whole genome sequencing will become the analysis method of choice. It is therefore important for ALSoD to be able to store next generation sequencing data and make the information readily accessible within the constraints of maintaining confidentiality when such detail exists and the data transfer implications of very large files. A noble goal will be to to make ALSoD a central repository for all ALS-related genome/exome sequencing. Careful consideration needs to be taken on setting up pipelines, the level of computing power required to carry out this enormous task and the type of files to use. Tools currently in development by the research community include PlinkSeq, which is an offshoot of the popular genetics analysis program Plink using a SQL server back end [611].

With the advent of newer technologies, international collaborations between scientists and greater awareness of research aims by those affected by ALS, what we are learning about ALS is expanding every day. We therefore aim to make the ALS Online genetics Database a one-stop shop integrating all genetic data in ALS.

More bioinformatics hyperlinks and integration to websites for the use of clinicians and non-clinicians, scientists and laymen are gradually being developed to serve both the scientific and non-scientific community by developing various sections on the website. Currently, the database serves only ALS researchers and clinicians but we hope to involve more with the patients through a collaborative effort with the PatientsLikeMe web team [612]. Plans are currently being carried out to extend the use of information on the database to ALS patients worldwide. Apart from the ability to search the database for information on mutations and the patients as published in journals, a direct link from a patient-based website like patientslikeme.com will be highly desirable. Modalities on the execution of this process are currently being discussed by both database teams. Links to and from both websites will allow patients currently registered on PatientsLikeMe website access to more information on their details about their diagnosis.

ALSoD collaborated with Dr Andrew Martin's bioinformatics group at UCL in 2008 but every effort to establish the broken link has proved abortive. So, I hope to either collaborate with another bioinformatics group who have mutant structures on SOD1 and SOD2 genes [613] or program the mutant structures for variants in ALSoD ourselves within the team. A positive response has been received from the Bioinfogroup team (http://bioinfogroup.com/database) and we hope to commence communication, planning and implementation in the nearest future.

A more advanced development of an integrated application that enhances collaboration among biologists and bioinformaticians [506] will be incorporated into ALSoD.

8.7.2 Larger Datasets

Genetic analyses of existing and future genetic studies is hereby automated, and presented in such a way as to allow meta-analysis in user-configurable ways, with follow-up of user-defined queries using a simple interface. In future, this will include larger populations for larger analysis especially with the introduction of improved technology on next generation sequencing.

8.7.3 Graphical displays

"A picture is worth a thousand words" is a common saying. The introduction of graphical display on ALSoD has increased the usability of the database. Further graphical representation of information proposed is to create a more detailed geographical display of epidemiology study data on Google Earth. Also, structural display of variants on genes will be developed for genes available in ALSoD.

8.7.4 Analysis

A publication utilised ALSoD dataset with other databases to evaluate the pathogenicity of ALS disease variants [555]. I hope to expand the credibility analysis of genes by including his formula as one of the criteria for ranking genes. I also hope to include information from animal models as one of the criteria.

Also, talks have commenced with one of my colleagues who is a Bioinformatician, Dr Kuang Lin of Neuroscience department. The lack of an established golden standard in the field of ALS is a major challenge. We hope to develop a 'silver standard' for ranking genes by selecting random controls which should help us to define the rules of detecting correlated pathogenic genes. We are trying to find the contrast between our controls and established genes. We will try to formulate a mathematical computational model to achieve this.

The publication on credibility score is drawing a lot of attention and so, in future, we hope to apply more criteria like prior information from animal models. Also, we hope to extend the functionalities to make it a template for other diseases that can provide these data as criteria.

8.7.5 Development

Feedback from users suggesting that extension of ALSoD could include integrated information from other species such as mouse and drosophila is welcomed. This was implemented by including publications on genes associated with animal models on the website and commencing the multiple alignments on SOD1 gene only. Once we have a critical mass of information we will seek collaboration with experts in each field to help with analysis and presentation of data.

An Android application was developed following a mobile website creation which allows users to switch between the full version and the mobile version. A further development on a wider platform will include a windows app and an iOS application (for iPhone and iPad).

Chapter 9 CONCLUSIONS

The original rationale for the ALSoD database in 1999 was as a central resource collating information on *SOD1* mutations so that researchers could data-mine for clinical/genetic correlations or other patterns in the data. This remains an important goal of the database. A few issues with the original design meant that registration and access was cumbersome, discouraging users from registering. In addition, the database relied completely on the goodwill of the research community for updates, there was no easy way to include new genes, and there was no easy way to incorporate the latest advances in bioinformatics. These issues have all been addressed in modernizing ALSoD over the last few years.

9.1 Summary for Chapter 3

Chapter 3 is the publication on “ALSoD: A User-Friendly Online Bioinformatics Tool for Amyotrophic Lateral Sclerosis Genetics”.

The ALSoD database has been transformed from a storage database into an analytical tool for ALS genetic data. It has metamorphosed from a single gene, locus-specific mutation storage facility to a multigene, disease-specific analytical database. Patterns in the curated genotype and phenotype information are discovered.

Feedback information received from users before the current development revealed that researchers who went on the database to utilize the resources were discouraged from the lengthy registration process required before access could be granted to use the database. There were also issues raised about the low frequency of data update.

It is difficult for a researcher to keep up with the vast volume of data regularly reported in the field of ALS globally. The role of ALSoD is to collate this prodigious amount of information into an easily comprehensible and convenient dataset.

9.2 Summary for Chapter 4

Chapter 4 is the publication on “Keeping up with genetic discoveries in amyotrophic lateral sclerosis: The ALSoD and ALSGene databases”.

Web resources (http://alsod.iop.kcl.ac.uk/) and (http://www.als-gene.org) were independently developed to provide a summary of up-to-date most interesting genetic discoveries in ALS resulting from the utilisation of these resources. A joint publication was written to explain the main focus of the databases, the different

approaches imbibed to reach a common goal and the uniqueness of each resource through collaboration. The main focus for me as a first joint author is ALSoD and how the development of this unique resource benefits the ALS research community worldwide.

A GWAS analysis from five different populations was developed and open access to registered users is provided to compare their data with the GWAS data. Codes and scripts were written in programming languages like JavaScript, C#, T-SQL, XML and VB.NET integrated under the ASP.NET platform. This enhances flexibility and manipulation by users of the database through the webpage.

This tool enlightens researchers on significant SNPs where larger data are available as one of the limitations of GWAS is the unavailability of larger data. The accessibility of several thousand well-phenotyped cases is a paramount aspect of any GWA study focussed at finding genetic factors with inconspicuous effects [614]. Hence, an on-the-fly analytical tool enlarging the dataset for better analysis of the ALS disease was established.

A recent GWAS study from eight independent studies (the largest Genome Wide collaboration in ALS so far) is available on http://alsod.iop.kcl.ac.uk/GWA2/gwas_fogh.aspx. This is the latest development to the database allowing users to query ALSoD for a specific SNP id and the result is a display of the SNP, genomic position, odds ratio and P-Value in a tabular form. The meta-analysis includes data from 13,225 individuals analyzed for 6,138,740 overlapping markers [544]. A graphical representation of the data in haploview form will be included on the webpage as part of the future work for this meta- analysis display in ALSoD.

ALSoD systematically links to other bioinformatics tools, scientific and non-scientific external comprehensive databases using unique identifiers which are stored on the gene table. Gene variants are mapped to Google Earth and embedded on the website for viewing. A user interacts with this application by zooming in and out, clicking on the yellow pins to show the name of the gene and its geographical location on the globe. Interactions between selected genes on ALSoD are explored on GeneMANIA interface which is an interactive collaboration with the bioinformatics tool.

ALSGene and ALSoD have joined forces by hyperlinking related information existing on the two independent databases to provide a unified and stress-free navigation to users.

9.3 Summary for Chapter 5

Chapter 5 is the publication on the “Credibility Analysis of Putative Disease-Causing Genes Using Bioinformatics”.

The focus of the credibility analysis is to weigh the evidences that a gene could be regarded as an ‘ALS gene’ in comparison to other genes available on the database. Credibility as defined in a previous association study utilised three criteria: amount of evidence, replicability of result and protection from bias [112].

Three bioinformatics tools freely available online were used to predict the pathogenicity of a variant in an ALS gene. These are PANTHER (Protein Analysis Through Evolutionary Relationships), SIFT (Sorting Intolerant From Tolerant) and PolyPhen (Polymorphism Phenotyping).

Mutation and patient data on 14 Familial ALS genes available in ALSoD where updated using publication search engines and other locus-specific databases. Nine criteria were embedded on the user-configurable page for researchers to select. Genes were ranked according to the dense ranking system where gaps are avoided and two compulsory criteria (Rank by mutation and Rank by patients) common to all the genes are automatically selected.

Using Surveymonkey tool, a survey was carried out on experts in the field of ALS who have publications as first or last author on ALS genetics. This was to protect the design from bias by examining responses from experts.

One of the downsides to the prediction of pathogenicity is in the use of Polyphen software which allows single entry of mutation on their website. So, on ALSoD, only mutations recorded have been analysed by the 3rd bioinformatics tool (polyphen). If new mutations are added by a user to the database, the Administrator requires a manual online generation of analysis on the polyphen website and then involves the update of the displayed result to the POLYPHEN_Score table.

Another issue is the C9orf72 gene which has been excluded from the analysis due to its varied repeat expansion nature. This makes it unadjustable to the structure of the credibility ranking despite the overwhelming amount of evidence from publications.

9.4 Summary for Chapter 6

Chapter 6 is the publication on the “Development of a Smartphone App for a Genetics Website: The Amyotrophic Lateral Sclerosis Online Genetics Data”.

A mobile version of the desktop version was created to allow more accessibility and easy readability of information on the ALSoD website on any device. A summarized version of the full desktop version was developed for mobile devices like a smartphone or tablet with an internet access.

Blogs are social networking media for presenting ideas or sharing ones thoughts around the globe. A blog was built on this website for both the desktop and mobile views to enable users of the database discuss hot topics or controversies in the field of ALS. Users are not obliged to sign on or log in before comments could be posted even though this constitutes an unusual risk unlike other blog sites. Also, without a registration process, users could pose sensitive questions or hot topics anonymously without fear or favour.

Analysis was carried out within 12 months to analyse the effect of introducing a mobile platform and through the in house-built monitoring system and the external Google Analytics tool, it was discovered that ALSoD has experienced a surge in the use of the website.

9.5 Conclusion

The main goal of genetics is to determine the genotypes that explain phenotypes. Due to the availability of inexpensive technology and instruments capable of producing millions of sequence data in genetics at an amazing speed, it is important that researchers are able to utilise the information provided [104]. ALSoD is therefore a valuable tool to aid researchers in accomplishing this goal.

Because the database summarizes ALS genetics including family and genome-wide association studies, the relevance of different findings can be quickly assessed and prioritized by researchers. Furthermore, the raw data can be more easily accessed because of the strong collaborative links with other groups and the inclusive nature of ALSoD.

Linkage and association studies, both of candidate genes and as part of genome-wide scans have generated many potential genetic findings of interest, and some of these are specific to certain geographies or ethnic groups [615]. The pace of new discoveries is rapid and the interpretation of findings, their relationship to existing phenotypes and gene discoveries, and the integration of new genetic data requires a coordinated, systematic and logical approach, which ALSoD provides.

Efforts are growing rapidly by the ALS-research community to complement the role of ALSoD around the world. Recently, an ALS mutation database constructed as part of the Life Science Integrated Database Project conducted by the Japan Ministry of Education, Culture, Sports, Science, and Technology was published. It contains their original experimental results and published data extracted from scientific journals. The database is expected to play a complementary role to the ALSoD database especially in collecting variations in the Asian region [485].

The ALSGene database has also emerged to “provide a comprehensive, unbiased and regularly updated field synopsis of genetic association studies performed in ALS. Once content creation is completed, one of its main features will consist of up-to-date meta-analyses for all eligible polymorphisms with sufficient data” [539]. A successfully implemented methodology earlier developed for meta-analyses in the field of Alzheimer [616], Multiple Sclerosis [617], Parkinson’s [618] and Schizophrenia [619] were extended to Amyotrophic Lateral Sclerosis [539].

Scripts and tables have been rewritten to remove redundant fields and converted to SQL from Cold Fusion, bringing the system up-to-date. The signup process has been simplified and a user feedback page added. Different levels of access are available for non-registered, registered and administrator users.

The internal database structure has been redesigned, so that it is now a relatively simple process for new genes or loci to be added with the same template of links and data available, and automatically integrated into the search and table schemes of the website.

We have reciprocal links to the HGVbaseG2P website at Leicester University Genetics Department, which is a genotype to phenotype database, so that visitors to either site can follow links to the other. ALSoD is also listed in the Online Mendelian Inheritance in Man (OMIM) database, again allowing visitors to either site access to the other.

Each gene page has specific hyperlinks to relevant bioinformatics sources. There are typically more than 20 such links for any gene, including the expected links to PubMed, HapMap, genome browsers at UCSC and Ensembl and OMIM, but also directly querying more unusual information sources such as Wolfram Alpha. Where possible, I have also mapped ALS-associated genetic variations geographically on Google Earth.

ALSoD automatically searches online for new information that is likely to be of interest to ALS genetics researchers, and alerts the development team for review. As a result it is easier to keep the database up-to-date. Some of this information is automatically made available before review through the ALS News link.

Genetic studies and findings on ALSoD are classified into those found by linkage of familial ALS genes, those found by candidate gene association, and the results of genome-wide association studies.

One of the goals of the project is to automatically collect and integrate data. This was partly achieved by automatically getting notifications of new ALS genes or mutations from journals but not without the human intervention of checking the publications manually to update the database with current relevant information. For a genetic database that requires accurate information like ALSoD, human intervention is inevitable.

ALSoD is truly a representative of the research community. It is an adopted project of the World Federation of Neurology and European ALS Consortium and currently funded by four different research charities, ALSA, MNDA, ALS Canada and MNDA Iceland. An average of 450 visitors uses the database daily from more than 150 countries. The most frequent users are from the USA, countries in EU, China and UK including some visits from countries in Africa and Asia. The web resource has been cited over 800 times outside the UK. 53% of visitors recorded monthly are new and more than 40 thousand unique visits have been recorded in the past 4 years.

Although there are other databases available online as discussed above, ALSoD is unique in being a genotype: phenotype database, disease specific and detailed regularly updated database by ALSoD curators and the research community around the world.

References

1. Goodreads.com. [cited 2013 4th November]; Available from: http://www.goodreads.com/quotes/28440-the-greatest-wealth-is-health.
2. Deeg, D.J., M.J. van Vliet, and J.W. Kardaun, *Understanding the Mortality Decline at Older Ages.* Annual Review of Gerontology and Geriatrics, Volume 33, 2013: Healthy Longevity, 2013. **33**: p. 261.
3. Hoeymans, N., et al., *The disabling effect of diseases: A study on trends in diseases, activity limitations, and their interrelationships.* Journal Information, 2012. **102**(1).
4. Wise, D.A., *Retirement against the demographic trend: More older people living longer, working less, and saving less?* Demography, 1997. **34**(1): p. 83-95.
5. Plus, M.D.-M. 2012 18 April 2012 [cited 4th November 2013]; Available from: http://www.nlm.nih.gov/medlineplus/mplusdictionary.html.
6. Roizen, N.J. and D. Patterson, *Down's syndrome.* The Lancet, 2003. **361**(9365): p. 1281-1289.
7. Verp, M.S., *Chromosomal Disorders in Pregnancy*, in *Principles of Medical Therapy in Pregnancy*. 1985, Springer. p. 1223-1229.
8. Niebuhr, E., *Down's syndrome.* Humangenetik, 1974. **21**(1): p. 99-101.
9. Salk, J. and D. Salk, *Control of influenza and poliomyelitis with killed virus vaccines.* Science, 1977. **195**(4281): p. 834-847.
10. Salk, J.E., *POLIOMYELITIS VACCINE PREPARATION AND ADMINISTRATION ANALYSIS OF BASIC PREMISES AND CURRENT PRACTICES.* Journal of the American Medical Association, 1959. **169**(16): p. 1829-1838.
11. McKusick, V.A., *A catalog of human genes and genetic disorders.* Mendelian inheritance in man, 1998: p. 226-227.
12. Risch, N. and K. Merikangas, *The future of genetic studies of complex human diseases.* Science-AAAS-Weekly Paper Edition, 1996. **273**(5281): p. 1516-1517.
13. Phukan, J., N. Pender, and O. Hardiman, *Cognitive impairment in amyotrophic lateral sclerosis.* The Lancet Neurology, 2007. **6**(11): p. 994-1003.
14. Veldink, J., L. Berg, and J. Wokke, *The future of motor neuron disease.* Journal of Neurology, 2004. **251**(4): p. 491-500.
15. Andersen, P.M. and A. Al-Chalabi, *Clinical genetics of amyotrophic lateral sclerosis: what do we really know?* Nature Reviews Neurology, 2011. **7**(11): p. 603-615.
16. Charcot, J. and A. Joffroy, *Deux cas d'atrophie musculaire progressive.* Arch Physiol Norm Pathol, 1869. **2**: p. 744-760.
17. Cleveland, D.W. and J.D. Rothstein, *From Charcot to Lou Gehrig: deciphering selective motor neuron death in ALS.* Nature Reviews Neuroscience, 2001. **2**(11): p. 806-819.
18. Turner, M.R. and K. Talbot, *14 Motor Neuron Disease.* Understanding Medical Research: The Studies That Shaped Medicine, 2012: p. 151.
19. Al-Chalabi, A., et al., *Recessive amyotrophic lateral sclerosis families with the D90A SOD1 mutation share a common founder: evidence for a linked protective factor.* Human Molecular Genetics, 1998. **7**(13): p. 2045.
20. Cummings, J.L., et al., *Defining and diagnosing involuntary emotional expression disorder.* CNS spectrums, 2006. **11**(6): p. 1.
21. Donkervoort, S. and T. Siddique, *Amyotrophic Lateral Sclerosis Overview.* GeneReviews. Seattle (WA): University of Washington, Seattle, 1993. **2001**.
22. Pasquali, L., et al., *The role of autophagy: what can be learned from the genetic forms of amyotrophic lateral sclerosis.* CNS &# 38; Neurological Disorders-Drug Targets (Formerly Current Drug Targets, 2010. **9**(3): p. 268-278.
23. Al-Chalabi, A. and O. Hardiman, *The epidemiology of ALS: a conspiracy of genes, environment and time.* Nature Reviews Neurology, 2013.
24. Ray, S.S. and P.T. Lansbury, *A possible therapeutic target for Lou Gehrig's disease.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(16): p. 5701-5702.
25. Byrne, S., et al., *Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis.* Journal of Neurology, Neurosurgery and Psychiatry, 2011. **82**(6): p. 623.
26. Valdmanis, P.N., et al., *Recent advances in the genetics of amyotrophic lateral sclerosis.* Curr Neurol Neurosci Rep, 2009. **9**(3): p. 198-205.
27. Van Damme, P. and W. Robberecht, *Recent advances in motor neuron disease.* Current Opinion in Neurology, 2009. **22**(5): p. 486-92.
28. van Es, M.A., et al., *Large-scale SOD1 mutation screening provides evidence for genetic heterogeneity in amyotrophic lateral sclerosis.* Journal of Neurology, Neurosurgery and Psychiatry, 2010. **81**(5): p. 562-6.

29. Wijesekera, L.C. and P.N. Leigh, *Amyotrophic lateral sclerosis.* Orphanet Journal of Rare Diseases, 2009. **4**(1): p. 3.
30. Dal Bello-Haas, V., A.D. Kloos, and H. Mitsumoto, *Physical therapy for a patient through six stages of amyotrophic lateral sclerosis.* Physical therapy, 1998. **78**(12): p. 1312-1324.
31. Pagon, R.A., et al., *Amyotrophic Lateral Sclerosis Overview.* 2001.
32. Worms, P.M., *The epidemiology of motor neuron diseases: a review of recent studies.* Journal of the Neurological Sciences, 2001. **191**(1-2): p. 3-9.
33. Andersen, P., et al., *EFNS task force on management of amyotrophic lateral sclerosis: guidelines for diagnosing and clinical care of patients and relatives.* European Journal of Neurology, 2005. **12**(12): p. 921-938.
34. Andersen, P.M., et al., *EFNS guidelines on the clinical management of amyotrophic lateral sclerosis (MALS)–revised report of an EFNS task force.* European Journal of Neurology, 2012. **19**(3): p. 360-375.
35. Brooks, B.R., *El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial" Clinical limits of amyotrophic lateral sclerosis" workshop contributors.* Journal of the Neurological Sciences, 1994. **124**: p. 96.
36. Brooks, B.R., et al., *El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis.* Amyotrophic lateral sclerosis and other motor neuron disorders: official publication of the World Federation of Neurology, Research Group on Motor Neuron Diseases, 2000. **1**(5): p. 293.
37. de Jong, S.W., et al., *Trial eligibility and the El Escorial criteria, the revised El Escorial criteria and the Awaji algorithm for ALS.* epidemiology, risk factors and treatment, 2012: p. 26.
38. Renton, A.E., et al., *A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD.* Neuron, 2011.
39. Neary, D., et al., *Frontotemporal lobar degeneration A consensus on clinical diagnostic criteria.* Neurology, 1998. **51**(6): p. 1546-1554.
40. Snowden, J.S., D. Neary, and D.M. MANN, *Frontotemporal dementia.* The British Journal of Psychiatry, 2002. **180**(2): p. 140-143.
41. Doble, A., *The pharmacology and mechanism of action of riluzole.* Neurology, 1996. **47**(6 Suppl 4): p. 233S-241S.
42. Miller, R., et al., *Clinical trials of riluzole in patients with ALS.* Neurology, 1996. **47**(4 Suppl 2): p. 86S-92S.
43. Miller, R., et al., *Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND).* Cochrane Database Syst Rev, 2007. **1**(1).
44. Armon, C., *Is the lithium-for-ALS genie back in the bottle? Not quite.* Neurology, 2010. **75**(7): p. 586-587.
45. Frost, J.H., et al. *How the Social Web Supports patient experimentation with a new therapy: The demand for patient-controlled and patient-centered informatics*. in *AMIA Annual Symposium Proceedings*. 2008. American Medical Informatics Association.
46. Wicks, P., et al., *Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm.* Nature biotechnology, 2011. **29**(5): p. 411-414.
47. Braunstein, M.L., *Empowering the Patient*, in *Health Informatics in the Cloud*. 2013, Springer. p. 67-79.
48. Chio, A., et al., *Lithium carbonate in amyotrophic lateral sclerosis Lack of efficacy in a dose-finding trial.* Neurology, 2010. **75**(7): p. 619-625.
49. Verstraete, E., et al., *Lithium lacks effect on survival in amyotrophic lateral sclerosis: a phase IIb randomised sequential trial.* Journal of Neurology, Neurosurgery & Psychiatry, 2012. **83**(5): p. 557-564.
50. Turner, M.R., M.J. Parton, and P.N. Leigh. *Clinical trials in ALS: an overview*. in *Seminars in neurology*. 2001. Copyright© 2001 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA. Tel.:+ 1 (212) 584-4662.
51. Atassi, N., et al., *Analysis of start-up, retention, and adherence in ALS clinical trials.* Neurology, 2013. **81**(15): p. 1350-1355.
52. Kong, Q., et al., *The Importance of Preclinical Trial Timing–a Potential Reason for the Disconnect between mouse Studies and Human Clinical Trials in ALS.* CNS Neuroscience & Therapeutics, 2012. **18**(9): p. 791-793.
53. Traynor, B., et al., *Neuroprotective agents for clinical trials in ALS A systematic assessment.* Neurology, 2006. **67**(1): p. 20-27.
54. Gladman, M., M. Cudkowicz, and L. Zinman, *Enhancing clinical trials in neurodegenerative disorders: lessons from amyotrophic lateral sclerosis.* Current opinion in neurology, 2012. **25**(6): p. 735-742.

55. Borasio, G., et al., *Standards of palliative care for patients with amyotrophic lateral sclerosis: results of a European survey.* Amyotrophic Lateral Sclerosis, 2001. **2**(3): p. 159-164.
56. Borasio, G.D. and R. Voltz, *Palliative care in amyotrophic lateral sclerosis.* Journal of Neurology, 1997. **244**(4): p. S11-S17.
57. Relief, W.E.C.o.C.P. and A.S. Care, *Cancer pain relief and palliative care: report of a WHO expert committee*. Vol. 804. 1990: World Health Organization.
58. Miller, R.G., et al., *Practice parameter: The care of the patient with amyotrophic lateral sclerosis (an evidence-based review) Report of the Quality Standards Subcommittee of the American Academy of Neurology.* Neurology, 1999. **52**(7): p. 1311-1311.
59. Leigh, P.N., et al., *Amyotrophic lateral sclerosis: a consensus viewpoint on designing and implementing a clinical trial.* Amyotrophic Lateral Sclerosis, 2004. **5**(2): p. 84-98.
60. Miller, R.G., et al., *Consensus guidelines for the design and implementation of clinical trials in ALS.* Journal of the neurological sciences, 1999. **169**(1): p. 2-12.
61. Brooks, B., et al., *The amyotrophic lateral sclerosis functional rating scale-Assessment of activities of daily living in patients with amyotrophic lateral sclerosis.* Archives of Neurology, 1996. **53**(2): p. 141-147.
62. Cedarbaum, J.M., et al., *The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function.* Journal of the neurological sciences, 1999. **169**(1): p. 13-21.
63. Carlborg, Ö. and C.S. Haley, *Epistasis: too often neglected in complex trait studies?* Nature Reviews Genetics, 2004. **5**(8): p. 618-625.
64. Miko, I., *Epistasis: Gene interaction and phenotype effects.* Nature Education, 2008. **1**(1).
65. Wang, Y., et al., *Genetic background affects epistatic interactions between two beneficial mutations.* Biology Letters, 2013. **9**(1).
66. Greene, C.S., et al., *Multifactor dimensionality reduction for graphics processing units enables genome-wide testing of epistasis in sporadic ALS.* Bioinformatics, 2010. **26**(5): p. 694-695.
67. Sreedharan, J. and R.H. Brown, *Amyotrophic lateral sclerosis: problems and prospects.* Annals of neurology, 2013. **74**(3): p. 309-316.
68. Migliore, L. and F. Coppedè, *Genetics, environmental factors and the emerging role of epigenetics in neurodegenerative diseases.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2009. **667**(1): p. 82-97.
69. Oates, N. and R. Pamphlett, *An epigenetic analysis of SOD1 and VEGF in ALS.* Amyotrophic Lateral Sclerosis, 2007. **8**(2): p. 83-86.
70. BMJ.com. *What is Epidemiology*. [cited 2013 7th December]; Available from: http://www.bmj.com/about-bmj/resources-readers/publications/epidemiology-uninitiated/1-what-epidemiology.
71. Beghi, E., *Are professional soccer players at higher risk for ALS?* Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2013. **14**(7-8): p. 501-506.
72. Belli, S. and N. Vanacore, *Proportionate mortality of Italian soccer players: is amyotrophic lateral sclerosis an occupational disease?* European journal of epidemiology, 2005. **20**(3): p. 237-242.
73. Chiò, A., et al., *ALS in Italian professional soccer players: the risk is still present and could be soccer-specific.* Amyotrophic lateral sclerosis, 2009. **10**(4): p. 205-209.
74. Horner, R., et al., *Occurrence of amyotrophic lateral sclerosis among Gulf War veterans.* Neurology, 2003. **61**(6): p. 742-749.
75. Tansey, C.M., P. Raina, and C. Wolfson, *Veterans' physical health.* Epidemiologic reviews, 2013. **35**(1): p. 66-74.
76. Weisskopf, M., et al., *Prospective study of military service and mortality from ALS.* Neurology, 2005. **64**(1): p. 32-37.
77. Al-Chalabi, A. and O. Hardiman, *The epidemiology of ALS: a conspiracy of genes, environment and time.* Nature Reviews Neurology, 2013. **9**(11): p. 617-628.
78. Beghi, E., et al., *The epidemiology and treatment of ALS: focus on the heterogeneity of the disease and critical appraisal of therapeutic trials.* Amyotrophic lateral sclerosis, 2011. **12**(1): p. 1-10.
79. Kurtzke, J.F. and G.W. Beebe, *Epidemiology of amyotrophic lateral sclerosis 1. A case-control comparison based on ALS deaths.* Neurology, 1980. **30**(5): p. 453-453.
80. Nelson, L., *Epidemiology of ALS.* Clinical neuroscience (New York, NY), 1995. **3**(6): p. 327.
81. Rettner, R. *Epigenetics: Definition & Examples*. 2013 [cited 2013 6th December]; Available from: http://www.livescience.com/37703-epigenetics.html.
82. choices, N. *MRI scan*. 2013 [cited 2013 12th November]; Available from: http://www.nhs.uk/conditions/mri-scan/pages/introduction.aspx.
83. Suhy, J., et al., *Early detection and longitudinal changes in amyotrophic lateral sclerosis by 1H MRSI.* Neurology, 2002. **58**(5): p. 773-779.
84. Bede, P., et al., *Grey matter correlates of clinical variables in amyotrophic lateral sclerosis (ALS): a neuroimaging study of ALS motor phenotype heterogeneity and cortical focality.* Journal of Neurology, Neurosurgery & Psychiatry, 2013. **84**(7): p. 766-773.

85. Filippi, M., et al., *Assessment of system dysfunction in the brain through MRI-based connectomics.* The Lancet Neurology, 2013. **12**(12): p. 1189-1199.
86. Lillo, P., et al., *Grey and white matter changes across the amyotrophic lateral sclerosis-frontotemporal dementia continuum.* PloS one, 2012. **7**(8): p. e43993.
87. Berthod, F. and F. Gros-Louis, *In Vivo and In Vitro Models to Study Amyotrophic Lateral Sclerosis.* Amyotrophic Lateral Sclerosis. Intech Rijeka, Croatia, 2012: p. 81-124.
88. Hewamadduma, C., et al., *160 Novel in vitro and in vivo models of ALS associated with TDP-43 proteinopathy.* Journal of Neurology, Neurosurgery & Psychiatry, 2012. **83**(3): p. e1-e1.
89. Bruijn, L., et al., *ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions.* Neuron, 1997. **18**(2): p. 327-338.
90. Gagliardi, S., et al., *Regulation of FMO and PON Detoxication Systems in ALS Human Tissues.* Neurotoxicity research, 2012: p. 1-8.
91. Rosen, D., et al., *Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis.* 1993.
92. Kwiatkowski, T.J., Jr., et al., *Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis.* Science, 2009. **323**(5918): p. 1205-8.
93. Vance, C., et al., *Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6.* Science, 2009. **323**(5918): p. 1208-11.
94. Sreedharan, J., et al., *TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis.* Science, 2008. **319**(5870): p. 1668-72.
95. Al-Chalabi, A., et al., *Deletions of the heavy neurofilament subunit tail in amyotrophic lateral sclerosis.* Hum Mol Genet, 1999. **8**(2): p. 157-64.
96. Greenway, M.J., et al., *ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis.* Nat Genet, 2006. **38**(4): p. 411-3.
97. Maruyama, H., et al., *Mutations of optineurin in amyotrophic lateral sclerosis.* Nature. **465**(7295): p. 223-6.
98. Mitchell, J., et al., *Familial amyotrophic lateral sclerosis is associated with a mutation in D-amino acid oxidase.* Proc Natl Acad Sci U S A. **107**(16): p. 7556-61.
99. Al-Chalabi, A., et al., *An estimate of amyotrophic lateral sclerosis heritability using twin data.* J Neurol Neurosurg Psychiatry, 2010. **In Press**.
100. Al-Chalabi, A. and C.M. Lewis, *Modelling the effects of penetrance and family size on rates of sporadic and familial disease.* Human Heredity, 2011. **71**(4): p. 281-288.
101. Cronin, S., et al., *Screening for replication of genome-wide SNP associations in sporadic ALS.* European Journal of Human Genetics, 2008. **17**(2): p. 213-218.
102. Davey, J.W., et al., *Genome-wide genetic marker discovery and genotyping using next-generation sequencing.* Nature Reviews Genetics, 2011. **12**(7): p. 499-510.
103. Metzker, M.L., *Sequencing technologies—the next generation.* Nature Reviews Genetics, 2009. **11**(1): p. 31-46.
104. Mardis, E., *The impact of next-generation sequencing technology on genetics.* Trends in Genetics, 2008. **24**(3): p. 133-141.
105. Bowser, R., M. Cudkowicz, and R. Kaddurah-Daouk, *Biomarkers for amyotrophic lateral sclerosis.* Expert Review of Molecular Diagnostics, 2006. **6**(3): p. 387-398.
106. Turner, M.R., et al., *Biomarkers in amyotrophic lateral sclerosis.* The Lancet Neurology, 2009. **8**(1): p. 94-109.
107. Iida, A., et al., *A functional variant in ZNF512B is associated with susceptibility to amyotrophic lateral sclerosis in Japanese.* Human Molecular Genetics, 2011. **20**(18): p. 3684-3692.
108. Andersen, P.M. and A. Al-Chalabi, *Clinical genetics of amyotrophic lateral sclerosis: what do we really know?* Nature Reviews Neurology, 2011.
109. Al-Chalabi, A., et al., *Genetic and epigenetic studies of amyotrophic lateral sclerosis.* Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2013. **14**(S1): p. 44-52.
110. Oyanagi, K., et al., *Amyotrophic lateral sclerosis of Guam: the nature of the neuropathological findings.* Acta neuropathologica, 1994. **88**(5): p. 405-412.
111. Cox, P.A. and O.W. Sacks, *Cycad neurotoxins, consumption of flying foxes, and ALS-PDC disease in Guam.* Neurology, 2002. **58**(6): p. 956-959.
112. Ioannidis, J., et al., *Assessment of cumulative evidence on genetic associations: interim guidelines.* International Journal of Epidemiology, 2008. **37**(1): p. 120.
113. Jones, A. and A.-C. Amrnar, *Genetics of the MNDs.* Amyotrophic Lateral Sclerosis and the Frontotemporal Dementias, 2012: p. 337.
114. Siddique, T., et al., *Linkage of a gene causing familial amyotrophic lateral sclerosis to chromosome 21 and evidence of genetic-locus heterogeneity.* New England Journal of Medicine, 1991. **324**(20): p. 1381-1384.
115. Ogasawara, M., et al., *Mild ALS in Japan associated with novel SOD mutation.* Nature Genetics, 1993. **5**(4): p. 323.

116. Deng, H.X., et al., *Amyotrophic lateral sclerosis and structural defects in Cu, Zn superoxide dismutase.* Science, 1993. **261**(5124): p. 1047.
117. Aoki, M., et al., *Familial amyotrophic lateral sclerosis (ALS) in Japan associated with H46R mutation in CuZn superoxide dismutase gene: A possible new subtype of familial ALS.* Journal of the neurological sciences, 1994. **126**(1): p. 77-83.
118. Elshafey, A., W.G. Lanyon, and J.M. Connor, *Identification of a new missense point mutation in exon 4 of the Cu/Zn superoxide dismutase (SOD-1) gene in a family with amyotrophic lateral sclerosis.* Human Molecular Genetics, 1994. **3**(2): p. 363.
119. Esteban, J., et al., *Identification of two novel mutations and a new polymorphism in the gene for Cu/Zn superoxide dismutase in patients with amyotrophic lateral sclerosis.* Human Molecular Genetics, 1994. **3**(6): p. 997.
120. Hirano, M., et al., *A new variant Cu/Zn superoxide dismutase (Val7--> Glu) deduced from lymphocyte mRNA sequences from Japanese patients with familial amyotrophic lateral sclerosis.* Biochemical and Biophysical Research Communications, 1994. **204**(2): p. 572-577.
121. Jones, C., et al., *Identification of a novel exon 4 SOD1 mutation in a sporadic amyotrophic lateral sclerosis patient.* Molecular and Cellular Probes, 1994. **8**(4): p. 329-330.
122. Jones, C.T., R.J. Swingler, and D.J.H. Brock, *Identification of a novel S0D1 mutation in an apparently sporadic amyotrophic lateral sclerosis patient and the detection of lle113Thr in three others.* Human Molecular Genetics, 1994. **3**(4): p. 649-650.
123. Kostrzewa, M., U. Burck-Lehmann, and U. MÃ¼ller, *Autosomal dominant amyotrophic lateral sclerosis: a novel mutation in the Cu/Zn superoxide dismutase-1 gene.* Human Molecular Genetics, 1994. **3**(12): p. 2261.
124. Nakano, R., et al., *A novel mutation in Cu/Zn superoxide dismutase gene in Japanese familial amyotrophic lateral sclerosis.* Biochemical and Biophysical Research Communications, 1994. **200**(2): p. 695-703.
125. Pramatarova, A., et al., *A two basepair deletion in the SOD 1 gene causes familial amyotrophic lateral sclerosis.* Human Molecular Genetics, 1994. **3**(11): p. 2061.
126. Rainero, I., et al., *SOD1 missense mutation in an Italian family with ALS.* Neurology, 1994. **44**(2): p. 347-347.
127. Siddique, T., H. Deng, and A. Hentati, *Identification of new mutations in familial amyotrophic lateral sclerosis.* American Journal of Human Genetics, 1994. **55**(CONF-941009--).
128. Suthers, G., et al., *Sporadic motoneuron disease due to familial SOD1 mutation with low penetrance.* Lancet, 1994. **344**(8939-0): p. 1773-1773.
129. Takahashi, H., et al., *Familial amyotrophic lateral sclerosis with a mutation in the Cu/Zn superoxide dismutase gene.* Acta Neuropathologica, 1994. **88**(2): p. 185-188.
130. Tsuda, T., et al., *Analysis of the functional effects of a mutation in SOD1 associated with familial amyotrophic lateral sclerosis.* Neuron, 1994. **13**(3): p. 727-736.
131. Andersen, P.M., et al., *Amyotrophic lateral sclerosis associated with homozygosity for an Asp90Ala mutation in CuZn-superoxide dismutase.* Nature Genetics, 1995. **10**(1): p. 61-66.
132. Aoki, M., et al., *Variance of age at onset in a Japanese family with amyotrophic lateral sclerosis associated with a novel Cu/Zn superoxide dismutase mutation.* Annals of Neurology, 1995. **37**(5): p. 676-679.
133. Deng, H.X., et al., *Two novel SOD1 mutations in patients with familial amyotrophic lateral sclerosis.* Human Molecular Genetics, 1995. **4**(6): p. 1113-1116.
134. Enayat, Z., et al., *Two novel mutations in the gene for copper zinc superoxide dismutase in UK families with amyotrophic lateral sclerosis.* Human Molecular Genetics, 1995. **4**(7): p. 1239.
135. Ikeda, M., et al., *A novel point mutation in the Cu/Zn superoxide dismutase gene in a patient with familial amyotrophic lateral sclerosis.* Human Molecular Genetics, 1995. **4**(3): p. 491-492.
136. Ikeda, M., et al., *Variable clinical symptoms in familial amyotrophic lateral sclerosis with a novel point mutation in the Cu/Zn superoxide dismutase gene.* Neurology, 1995. **45**(11): p. 2038-2042.
137. Orrell, R., et al., *A novel SOD mutant and ALS.* Nature, 1995. **374**(6522): p. 504.
138. Orrell, R., et al., *Familial amyotrophic lateral sclerosis with a point mutation of SOD-1: intrafamilial heterogeneity of disease duration associated with neurofibrillary tangles.* Journal of Neurology, Neurosurgery and Psychiatry, 1995. **59**(3): p. 266.
139. Pramatarova, A., et al., *Identification of new mutations in the Cu/Zn superoxide dismutase gene of patients with familial amyotrophic lateral sclerosis.* American Journal of Human Genetics, 1995. **56**(3): p. 592.
140. Sapp, P., et al., *Identification of three novel mutations in the gene for Cu/Zn superoxide dismutase in patients with familial amyotrophic lateral sclerosis.* Neuromuscular Disorders, 1995. **5**(5): p. 353-357.
141. Sjalander, A., et al., *The D90A mutation results in a polymorphism of Cu, Zn superoxide dismutase that is prevalent in northern Sweden and Finland.* Human Molecular Genetics, 1995. **4**(6): p. 1105-1108.

142. Yulug, I.G., et al., *An improved protocol for the analysis of SOD1 gene mutations, and a new mutation in exon 4.* Human Molecular Genetics, 1995. **4**(6): p. 1101-1104.
143. Abe, K., et al., *Clinical characteristics of familial amyotrophic lateral sclerosis with CuZn superoxide dismutase gene mutations.* Journal of the Neurological Sciences, 1996. **136**(1): p. 108-116.
144. Andersen, P., et al., *Autosomal recessive adult-onset amyotrophic lateral sclerosis associated with homozygosity for Asp90Ala CuZn-superoxide dismutase mutation.* Brain, 1996. **119**(4): p. 1153.
145. Hosler, B.A., et al., *Three novel mutations and two variants in the gene for Cu/Zn superoxide dismutase in familial amyotrophic lateral sclerosis.* Neuromuscular Disorders, 1996. **6**(5): p. 361-366.
146. Ince, P.G., et al., *Familial amyotrophic lateral sclerosis with a mutation in exon 4 of the Cu/Zn superoxide dismutase gene: pathological and immunocytochemical changes.* Acta Neuropathologica, 1996. **92**(4): p. 395-403.
147. Kostrzewa, M., M.S. Damian, and U. MÃ¼ller, *Superoxide dismutase 1: identification of a novel mutation in a case of familial amyotrophic lateral sclerosis.* Human Genetics, 1996. **98**(1): p. 48-50.
148. Morita, M., et al., *A novel two-base mutation in the Cu/Zn superoxide dismutase gene associated with familial amyotrophic lateral sclerosis in Japan.* Neuroscience Letters, 1996. **205**(2): p. 79-82.
149. Orrell, R.W., et al., *Difficulties in distinguishing sporadic from familial amyotrophic lateral sclerosis.* Annals of Neurology, 1996. **39**(6): p. 810-812.
150. Siddique, T., D. Nijhawan, and A. Hentati, *Molecular genetic basis of familial ALS.* Neurology, 1996. **47**(4 Suppl 2): p. 27S-35S.
151. Andersen, P., et al., *Phenotypic heterogeneity in motor neuron disease patients with CuZn-superoxide dismutase mutations in Scandinavia.* Brain, 1997. **120**(10): p. 1723.
152. Bereznai, B., et al., *A novel SOD1 mutation in an Austrian family with amyotrophic lateral sclerosis.* Neuromuscular Disorders, 1997. **7**(2): p. 113-116.
153. Cudkowicz, M., et al., *Epidemiology of mutations in superoxide dismutase in amyotrophic lateal sclerosis.* Annals of Neurology, 1997. **41**(2): p. 210-221.
154. Jackson, M., et al., *Copper/zinc superoxide dismutase 1 and sporadic amyotrophic lateral sclerosis: Analysis of 155 cases and identification of novel insertion mutation.* Annals of Neurology, 1997. **42**(5): p. 803-807.
155. Kawamata, J., et al., *Novel G16S (GGC-AGC) mutation in the SOD-1 gene in a patient with apparently sporadic young-onset amyotrophic lateral sclerosis.* Human Mutation, 1997. **9**(4): p. 356-358.
156. Kawata, A., et al., *Prominent sensory and autonomic disturbances in familial amyotrophic lateral sclerosis with a Gly93Ser mutation in the SOD1 gene.* Journal of the Neurological Sciences, 1997. **153**(1): p. 82-85.
157. Kikugawa, K., et al., *A missense mutation in the SOD1 gene in patients with amyotrophic lateral sclerosis from the Kii Peninsula and its vicinity, Japan.* Neurogenetics, 1997. **1**(2): p. 113-115.
158. Maeda, T., et al., *On intra-familial clinical diversities of a familial amyotrophic lateral sclerosis with a point mutation of Cu/Zn superoxide dismutase (Asn 86-Ser].* Nō to shinkei= Brain and nerve, 1997. **49**(9): p. 847.
159. Orrell, R., et al., *A novel mutation of SOD-1 (Gly 108 Val) in familial amyotrophic lateral sclerosis.* European Journal of Neurology, 1997. **4**(1): p. 48-51.
160. Orrell, R., et al., *Clinical and functional investigation of 10 missense mutations and a novel frameshift insertion mutation of the gene for copper-zinc superoxide dismutase in UK families with amyotrophic lateral sclerosis.* Neurology, 1997. **48**(3): p. 746-751.
161. Shaw, P.J., et al., *CNS tissue Cu/Zn superoxide dismutase (SOD1) mutations in motor neurone disease (MND).* Neuroreport, 1997. **8**(18): p. 3923.
162. Shaw, C., et al., *Familial amyotrophic lateral sclerosis. Molecular pathology of a patient with a SOD1 mutation.* Neurology, 1997. **49**(6): p. 1612-1616.
163. Watanabe, M., et al., *A novel missense point mutation (S134N) of the Cu/Zn superoxide dismutase gene in a patient with familal motor neuron disease.* Human Mutation, 1997. **9**(1): p. 69-71.
164. Zu, J.S., et al., *Exon 5 encoded domain is not required for the toxic function of mutant SOD1 but essential for the dismutase activity: identification and characterization of two new SOD1 mutations associated with familial amyotrophic lateral sclerosis.* Neurogenetics, 1997. **1**(1): p. 65-71.
165. Boukaftane, Y., et al., *Identification of six novel SOD1 gene mutations in familial amyotrophic lateral sclerosis.* Canadian Journal of Neurological Sciences, 1998. **25**: p. 192-196.
166. Hayward, C., et al., *Homozygosity for Asn86Ser mutation in the CuZn-superoxide dismutase gene produces a severe clinical phenotype in a juvenile onset case of familial amyotrophic lateral sclerosis.* Journal of Medical Genetics, 1998. **35**(2): p. 174.
167. Morita, M., et al., *A novel mutation Asp90Val in the SOD1 gene associated with Japanese familial ALS.* European Journal of Neurology, 1998. **5**(4): p. 389-392.

168. Nakanishi, T., et al., *Simple and defined method to detect the SOD-1 mutants from patients with familial amyotrophic lateral sclerosis by mass spectrometry.* Journal of Neuroscience Methods, 1998. **81**(1-2): p. 41-44.
169. Shaw, C., et al., *Mutations in all five exons of SOD-1 may cause ALS.* Annals of Neurology, 1998. **43**(3): p. 390-394.
170. Cudkowicz, M., et al., *Limited corticospinal tract involvement in amyotrophic lateral sclerosis subjects with the A4V mutation in the copper/zinc superoxide dismutase gene.* Annals of Neurology, 1998. **43**(6): p. 703-710.
171. Aguirre, T., et al., *Mutational analysis of the Cu/Zn superoxide dismutase gene in 23 familial and 69 sporadic cases of amyotrophic lateral sclerosis in Belgium.* European journal of human genetics: EJHG, 1999. **7**(5): p. 599.
172. Ceroni, M., et al., *Clustering of ALS patients in central Italy due to the occurrence of the L84F SOD1 gene mutation.* Neurology, 1999. **53**(5): p. 1064-1071.
173. Kohno, S., et al., *A novel mutation (Cys6Gly) in the Cu/Zn superoxide dismutase gene associated with rapidly progressive familial amyotrophic lateral sclerosis.* Neuroscience Letters, 1999. **276**(2): p. 135-137.
174. Orrell, R.W., et al., *Clinical characteristics of SOD1 gene mutations in UK families with ALS.* Journal of the Neurological Sciences, 1999. **169**(1-2): p. 56-60.
175. Penco, S., et al., *A SOD1 gene mutation in a patient with slowly progressing familial ALS.* Neurology, 1999. **53**(2): p. 404-404.
176. Cervenakova, L., et al., *Progressive muscular atrophy variant of familial amyotrophic lateral sclerosis (PMA/ALS).* Journal of the Neurological Sciences, 2000. **177**(2): p. 124-130.
177. Shimizu, T., et al., *Autonomic failure in ALS with a novel SOD1 gene mutation.* Neurology, 2000. **54**(7): p. 1534-1537.
178. Slominsky, P.A., et al., *CuZn superoxide dismutase gene in sporadic amyotrophic lateral sclerosis patients from Russia: Asp90Ala (D90A) mutation and novel rare polymorphism IVS3+ 35 A> C.* Human Mutation, 2000. **16**(3): p. 277-278.
179. Weber, M., et al., *Preserved slow conducting corticomotoneuronal projections in amyotrophic lateral sclerosis with autosomal recessive D90A CuZn-superoxide dismutase mutation.* Brain, 2000. **123**(7): p. 1505.
180. Winter, S.M., et al., *Recessively inherited amyotrophic lateral sclerosis: a German family with the D90A CuZn-SOD mutation.* Journal of Neurology, 2000. **247**(10): p. 783-786.
181. Gellera, C., *Genetics of ALS in Italian families.* Amyotrophic Lateral Sclerosis, 2001. **2**(1): p. 43-46.
182. Hand, C.K., et al., *Compound heterozygous D90A and D96N SOD1 mutations in a recessive amyotrophic lateral sclerosis family.* Annals of Neurology, 2001. **49**(2): p. 267-271.
183. Kato, M., et al., *Marked reduction of the Cu/Zn superoxide dismutase polypeptide in a case of familial amyotrophic lateral sclerosis with the homozygous mutation.* Neuroscience Letters, 2001. **312**(3): p. 165-168.
184. Mase, G., et al., *ALS with variable phenotypes in a six-generation family caused by leu144phe mutation in the SOD1 gene.* Journal of the Neurological Sciences, 2001. **191**(1): p. 11-18.
185. Murakami, T., et al., *A novel SOD1 gene mutation in familial ALS with low penetrance in females.* Journal of the Neurological Sciences, 2001. **189**(1-2): p. 45-47.
186. Takehisa, Y., et al., *Familial amyotrophic lateral sclerosis with a novel Leu126Ser mutation in the copper/zinc superoxide dismutase gene showing mild clinical features and Lewy body-like hyaline inclusions.* Archives of Neurology, 2001. **58**(5): p. 736.
187. Alexander, M.D., et al., *"True" sporadic ALS associated with a novel SODâ€• 1 mutation.* Annals of Neurology, 2002. **52**(5): p. 680-683.
188. Garcia-Redondo, A., et al., *Molecular analysis of the superoxide dismutase 1 gene in Spanish patients with sporadic or familial amyotrophic lateral sclerosis.* Muscle and Nerve, 2002. **26**(2): p. 274-278.
189. Iwai, K., et al., *Anticipation in familial amyotrophic lateral sclerosis with SOD1-G93S mutation.* Journal of Neurology, Neurosurgery and Psychiatry, 2002. **72**(6): p. 819.
190. Jonsson, P.A., et al., *CuZn-superoxide dismutase in D90A heterozygotes from recessive and dominant ALS pedigrees.* Neurobiology of Disease, 2002. **10**(3): p. 327-333.
191. Naini, A., et al., *Identification of a novel mutation in Cu/Zn superoxide dismutase gene associated with familial amyotrophic lateral sclerosis.* Journal of the Neurological Sciences, 2002. **198**(1-2): p. 17-19.
192. Segovia-Silvestre, T., et al., *A novel exon 3 mutation (D76V) in the SOD1 gene associated with slowly progressive ALS.* Amyotrophic Lateral Sclerosis, 2002. **3**(2): p. 69-74.
193. Aksoy, H., et al., *A4T mutation in the SOD1 gene causing familial amyotrophic lateral sclerosis.* Neuroepidemiology, 2003. **22**: p. 235-238.

194. Andersen, P., et al., *Sixteen novel mutations in the Cu/Zn superoxide dismutase gene in amyotrophic lateral sclerosis: a decade of discoveries, defects and disputes.* Amyotrophic Lateral Sclerosis, 2003. **4**(2): p. 62-73.
195. Ferrera, L., et al., *An Italian dominant FALS Leu144Phe SOD1 mutation: genotype-phenotype correlation.* Amyotrophic Lateral Sclerosis, 2003. **4**(3): p. 167-170.
196. Kim, N.H., et al., *A novel SOD1 gene mutation in a Korean family with amyotrophic lateral sclerosis.* Journal of the Neurological Sciences, 2003. **206**(1): p. 65-69.
197. Mayeux, V., et al., *N19S, a new SOD1 mutation in sporadic amyotrophic lateral sclerosis: no evidence for disease causation.* Annals of Neurology, 2003. **53**(6): p. 815-818.
198. Rezania, K., et al., *A rare Cu/Zn superoxide dismutase mutation causing familial amyotrophic lateral sclerosis with variable age of onset, incomplete penetrance and a sensory neuropathy.* Amyotrophic Lateral Sclerosis, 2003. **4**(3): p. 162-166.
199. Niemann, S., et al., *Familial ALS in Germany: origin of the R115G SOD1 mutation by a founder effect.* Journal of Neurology, Neurosurgery and Psychiatry, 2004. **75**(8): p. 1186.
200. Nogales-Gadea, G., et al., *A novel exon 5 mutation (N139H) in the SOD1 gene in a Spanish family associated with incomplete penetrance.* Journal of the Neurological Sciences, 2004. **219**(1-2): p. 1-6.
201. Ohi, T., et al., *Familial amyotrophic lateral sclerosis with His46Arg mutation in Cu/Zn superoxide dismutase presenting characteristic clinical features and Lewy body-like hyaline inclusions.* Journal of the Neurological Sciences, 2004. **225**(1-2): p. 19-25.
202. Sato, T., et al., *Identification of two novel mutations in the Cu/Zn superoxide dismutase gene with familial amyotrophic lateral sclerosis: mass spectrometric and genomic analyses.* Journal of the Neurological Sciences, 2004. **218**(1-2): p. 79-83.
203. Shi, S., et al., *Identification of the mutation of SOD1 gene in a familial amyotrophic lateral sclerosis].* Zhonghua yi xue yi chuan xue za zhi= Zhonghua yixue yichuanxue zazhi= Chinese journal of medical genetics, 2004. **21**(2): p. 149.
204. Tan, C.F., et al., *Familial amyotrophic lateral sclerosis with bulbar onset and a novel Asp101Tyr Cu/Zn superoxide dismutase gene mutation.* Acta Neuropathologica, 2004. **108**(4): p. 332-336.
205. Antonyuk, S., et al., *Structural consequences of the familial amyotrophic lateral sclerosis SOD1 mutant His46Arg.* Protein Science, 2005. **14**(5): p. 1201-1213.
206. Battistini, S., et al., *SOD1 mutations in amyotrophic lateral sclerosis.* Journal of Neurology, 2005. **252**(7): p. 782-788.
207. Sato, T., et al., *Rapid disease progression correlates with instability of mutant SOD1 in familial ALS.* Neurology, 2005. **65**(12): p. 1954-1957.
208. Zhang, H., et al., *A rare Cu/Zn superoxide dismutase mutation causing familial amyotrophic lateral sclerosis with variable age of onset and incomplete penetrance in China.* Amyotrophic Lateral Sclerosis, 2005. **6**(4): p. 234-238.
209. Andersen, P., *Amyotrophic lateral sclerosis genetics with Mendelian inheritance.* Brown RH Jr, Swash M, Pasinelli PAmyotrophic lateral sclerosis. Lodon: Informa Healthcare, 2006: p. 187-207.
210. Corrado, L., et al., *SOD1 gene mutations in Italian patients with sporadic amyotrophic lateral sclerosis (ALS).* Neuromuscular Disorders, 2006. **16**(11): p. 800-804.
211. Fong, G.C.Y., et al., *Clinical phenotypes of a large Chinese multigenerational kindred with autosomal dominant familial ALS due to Ile149Thr SOD1 gene mutation.* Amyotrophic Lateral Sclerosis, 2006. **7**(3): p. 142-149.
212. Gamez, J., et al., *Mutational analysis of the Cu/Zn superoxide dismutase gene in a Catalan ALS population: Should all sporadic ALS cases also be screened for SOD1?* Journal of the Neurological Sciences, 2006. **247**(1): p. 21-28.
213. Li, X., et al., *P142 SCREENING OF SUPEROXIDE DISMUTASE 1 GENE MUTATIONS IN CHINESE PATIENTS WITH FAMILIAL AND SPORADIC AMYOTROPHIC LATERAL SCLEROSIS.* Amyotrophic Lateral Sclerosis., 2006. **(Suppl 1)**(7): p. 147-148.
214. Regal, L., et al., *The G93C mutation in superoxide dismutase 1: clinicopathologic phenotype and prognosis.* Archives of Neurology, 2006. **63**(2): p. 262.
215. Stewart, H., et al., *Clinicopathological phenotype of ALS with a novel G72C SOD1 gene mutation mimicking a myopathy.* Muscle and Nerve, 2006. **33**(5): p. 701-706.
216. Beck, M., M. Sendtner, and K.V. Toyka, *Novel SOD1 N86K mutation is associated with a severe phenotype in familial ALS.* Muscle and Nerve, 2007. **36**(1): p. 111-114.
217. Holmoy, T., K. Bjorgo, and P.M. Roos, *Slowly progressing amyotrophic lateral sclerosis caused by H46R SOD1 mutation.* European Neurology, 2007. **58**(1): p. 57-58.
218. Kim, H.Y., et al., *Clinical characteristics of familial amyotrophic lateral sclerosis with a Phe20Cys mutation in the SOD1 gene in a Korean family.* Amyotrophic Lateral Sclerosis, 2007. **8**(2): p. 73-78.
219. Kim, W., et al., *Anticipation and phenotypic heterogeneity in Korean familial amyotrophic lateral sclerosis with superoxide dismutase 1 gene mutation.* Journal of Clinical Neurology (Seoul, Korea), 2007. **3**(1): p. 38.

220. Naini, A., et al., *Identification of a novel D109Y mutation in Cu/Zn superoxide dismutase (sod1) gene associated with amyotrophic lateral sclerosis.* Journal of the Neurological Sciences, 2007. **254**(1-2): p. 17-21.
221. Broom, W.J., et al., *50bp deletion in the promoter for superoxide dismutase 1 (SOD1) reduces SOD1 expression in vitro and may correlate with increased age of onset of sporadic amyotrophic lateral sclerosis.* Amyotrophic Lateral Sclerosis, 2008. **9**(4): p. 229-237.
222. Corti, S., et al., *Amyotrophic lateral sclerosis linked to a novel< i> SOD1</i> mutation with muscle mitochondrial dysfunction.* Journal of the neurological sciences, 2009. **276**(1): p. 170-174.
223. Eisen, A., et al., *SOD1 gene mutations in ALS patients from British Columbia, Canada: clinical features, neurophysiology and ethical issues in management.* Amyotrophic Lateral Sclerosis, 2008. **9**(2): p. 108-119.
224. Suzuki, M., et al., *Familial amyotrophic lateral sclerosis with Gly93Ser mutation in Cu/Zn superoxide dismutase: a clinical and neuropathological study.* Journal of the Neurological Sciences, 2008. **268**(1-2): p. 140-144.
225. Takahashi, Y., et al., *Development of a high-throughput microarray-based resequencing system for neurological disorders and its application to molecular genetics of amyotrophic lateral sclerosis.* Archives of Neurology, 2008. **65**(10): p. 1326.
226. Syriani, E., M. Morales, and J. Gamez, *The p. E22G mutation in the Cu/Zn superoxide-dismutase gene predicts a long survival time:: Clinical and genetic characterization of a seven-generation ALS1 Spanish pedigree.* Journal of the Neurological Sciences, 2009. **285**(1-2): p. 46-53.
227. Winkler, D.D., et al., *Structural and Biophysical Properties of the Pathogenic SOD1 Variant H46R/H48Qâ€* Biochemistry, 2009. **48**(15): p. 3436-3447.
228. Zinman, L., et al., *A mechanism for low penetrance in an ALS family with a novel SOD1 deletion.* Neurology, 2009. **72**(13): p. 1153.
229. Battistini, S., et al., *Severe familial ALS with a novel exon 4 mutation (L106F) in the SOD1 gene.* Journal of the Neurological Sciences, 2010. **293**(1): p. 112-115.
230. Felbecker, A., et al., *Four familial ALS pedigrees discordant for two SOD1 mutations: are all SOD1 mutations pathogenic?* Journal of Neurology, Neurosurgery and Psychiatry, 2010. **81**(5): p. 572-577.
231. Georgoulopoulou, E., et al., *A novel SOD1 mutation in a young amyotrophic lateral sclerosis patient with a very slowly progressive clinical course.* Muscle and Nerve, 2010. **42**(4): p. 596-597.
232. Giannini, F., et al., *D90A-SOD1 mutation in ALS: The first report of heterozygous Italian patients and unusual findings.* Amyotrophic Lateral Sclerosis, 2010. **11**(1-2): p. 216-219.
233. Holmoy, T., et al., *G127R: A novel SOD1 mutation associated with rapidly evolving ALS and severe pain syndrome.* Amyotrophic Lateral Sclerosis, 2010. **11**(5): p. 478-480.
234. Kobayashi, Z., et al., *FALS with Gly72Ser mutation in SOD1 gene: Report of a family including the first autopsy case.* Journal of the Neurological Sciences, 2010.
235. Laaksovirta, H., et al., *Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study.* The Lancet Neurology, 2010.
236. Millecamps, S., et al., *SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: genotype-phenotype correlations.* Journal of Medical Genetics, 2010. **47**(8): p. 554-560.
237. Origone, P., et al., *Enlarging clinical spectrum of FALS with TARDBP gene mutations: S393L variant in an Italian family showing phenotypic variability and relevance for genetic counselling.* Amyotrophic Lateral Sclerosis, 2010. **11**(1-2): p. 223-227.
238. Rabe, M., et al., *The epidemiology of CuZn-SOD mutations in Germany: a study of 217 families.* Journal of Neurology, 2010. **257**(8): p. 1298-1302.
239. Ricci, C., et al., *A novel exon 1 mutation (G10R) in the SOD1 gene in a patient with familial ALS.* Amyotrophic Lateral Sclerosis, 2010. **11**(5): p. 481-485.
240. Synofzik, M., et al., *The human G93A SOD1 phenotype closely resembles sporadic amyotrophic lateral sclerosis.* Journal of Neurology, Neurosurgery and Psychiatry, 2010. **81**(7): p. 764-767.
241. Takazawa, T., et al., *Familial Amyotrophic Lateral Sclerosis with a Novel G85S Mutation of Superoxide Dismutase 1 Gene: Clinical Features of Lower Motor Neuron Disease.* Internal Medicine, 2010. **49**(2): p. 183-186.
242. Tsai, C.P., et al., *FUS, TARDBP, and SOD1 mutations in a Taiwanese cohort with familial ALS.* Neurobiology of Aging, 2010. **32**(3): p. 553. e13-553. e21.
243. Visani, M., et al., *A novel T137A SOD1 mutation in an Italian family with two subjects affected by amyotrophic lateral sclerosis.* Amyotrophic Lateral Sclerosis, 2011(0): p. 1-4.
244. Baek, W., et al., *A novel codon4 mutation (A4F) in the SOD1gene in familial amyotrophic lateral sclerosis.* Journal of the Neurological Sciences, 2011.
245. Brotherton, T., et al., *A novel ALS SOD1 C6S mutation with implications for aggregation related toxicity and genetic counseling.* Amyotrophic Lateral Sclerosis, 2011(0): p. 1-5.

246. Conforti, F.L., et al., *Sporadic motor neuron disease in a familial novel SOD1 mutation: Incomplete penetrance or chance association?* Amyotrophic Lateral Sclerosis, 2011. **12**(3): p. 220-222.
247. del Grande, A., et al., *A novel L67P SOD1 mutation in an Italian ALS patient.* Amyotrophic Lateral Sclerosis, 2011(0): p. 1-3.
248. Hermann, A., et al., *Vocal cord paralysis and rapid progressive motor neuron disease by the I113F mutation in SOD1 gene.* Amyotrophic Lateral Sclerosis, 2011. **12**(5): p. 382-384.
249. Vela, A., et al., *SOD1-N196 mutation in a family with amyotrophic lateral sclerosis.* Neurologia (Barcelona, Spain), 2011.
250. Zhao, G., et al., *Clinical Features and Cu/Zn Superoxide Dismutase Gene Mutations in Two Mainland Chinese Families With Amyotrophic Lateral Sclerosis.* International Journal of Neuroscience, 2011. **121**(4): p. 191-195.
251. Berdynski, M., et al., *Recurrent G41S mutation in Cu/Zn superoxide dismutase gene (SOD1) causing familial amyotrophic lateral sclerosis in a large Polish family.* Amyotrophic Lateral Sclerosis, 2012. **13**(1): p. 132-136.
252. Brown, J.A., et al., *SOD1, ANG, TARDBP and FUS mutations in amyotrophic lateral sclerosis: A United States clinical testing lab experience.* Amyotrophic Lateral Sclerosis, 2012. **13**(2): p. 217-22.
253. Hu, J., et al., *A novel SOD1 mutation in amyotrophic lateral sclerosis with a distinct clinical phenotype.* Amyotrophic Lateral Sclerosis, 2012. **13**(1): p. 149-154.
254. Keckarevic, D., et al., *A novel P66S mutation in exon 3 of the SOD1 gene with early onset and rapid progression.* Amyotrophic Lateral Sclerosis, 2012(0): p. 1-4.
255. Origone, P., et al., *Fast course ALS presenting with vocal cord paralysis: Clinical features, bioinformatic and modelling analysis of the novel SOD1 Gly147Ser mutation.* Amyotrophic Lateral Sclerosis, 2012(0): p. 1-5.
256. Solski, J.A., et al., *A novel TARDBP insertion/deletion mutation in the flail arm variant of amyotrophic lateral sclerosis.* Amyotrophic Lateral Sclerosis, 2012. **13**(5): p. 465-470.
257. Diard-Detoeuf, C., et al., *Association of a paraneoplastic motor neuron disease with anti-Ri antibodies and a novel< i> SOD1 I18del</i> mutation.* Journal of the neurological sciences, 2013.
258. Dangoumau, A., et al., *A novel< i> SOD1</i> mutation p. V31A identified with a slowly progressive form of amyotrophic lateral sclerosis.* Neurobiology of aging, 2013. **35**(1): p. 266. e1-266. e4.
259. Nakamura, S., et al., *An autopsy case of sporadic amyotrophic lateral sclerosis associated with the I113T SOD1 mutation.* Neuropathology, 2013.
260. Naruse, H., et al., *Familial amyotrophic lateral sclerosis with novel A4D SOD1 mutation with late age at onset and rapid progressive course.* Neurology and Clinical Neuroscience, 2013. **1**(1): p. 45-47.
261. Klein, C.J., et al., *Novel SOD1 mutation discovered in atypical ALS by whole exome sequencing.* Journal of Neurology, Neurosurgery & Psychiatry, 2013.
262. Wang, Z., et al., *Identification of a novel missense (C7W) mutation of< i> SOD1</i> in a large familial amyotrophic lateral sclerosis pedigree.* Neurobiology of aging, 2013.
263. Cui, F., et al., *New mutation in the SOD1 (copper/zinc superoxide dismutase-1) gene in a Chinese amyotrophic lateral sclerosis (ALS) patient.* Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2013. **14**(7-8): p. 635-637.
264. Trojsi, F., et al., *an Italian kindred with FaLs due to c. 149t> C mutation in the soD1 gene: case report of an affected family member.* Acta Myologica, 2013. **32**(1): p. 23.
265. Kuzma-Kozakiewicz, M., et al., *Recurrent K3E mutation in Cu/Zn superoxide dismutase gene associated with amyotrophic lateral sclerosis.* Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2013. **14**(7-8): p. 608-614.
266. Otomo, A., et al., *Defective relocalization of ALS2/alsin missense mutants to Rac1-induced macropinosomes accounts for loss of their cellular function and leads to disturbed amphisome formation.* FEBS letters, 2011. **585**(5): p. 730-736.
267. Refsum, S. and S. Skillicorn, *Amyotrophic familial spastic paraplegia.* Neurology, 1954. **4**(1): p. 40.
268. HAMIDA, M.B.E.N., F. HENTATI, and C.B.E.N. HAMIDA, *Hereditary motor system diseases (chronic juvenile amyotrophic lateral sclerosis).* Brain, 1990. **113**(2): p. 347.
269. Gragg, G., M. Fogelson, and R. Zwirecki, *Juvenile amyotrophic lateral sclerosis in two brothers from an inbred community.* Birth Defects Original Article Series, 1971. **7**(1): p. 222.
270. Silver, J., *Familial spastic paraplegia with amyotrophy of the hands.* Journal of Neurology, Neurosurgery and Psychiatry, 1966. **29**(2): p. 135.
271. Hentati, A., et al., *Linkage of recessive familial amyotrophic lateral sclerosis to chromosome 2q33â€“q35.* Nature Genetics, 1994. **7**(3): p. 425-428.
272. Hentati, A., et al., *The gene locus for one form of juvenile amyotrophic lateral sclerosis maps to chromosome 2. (Abstract).* Am. J. Hum. Genet., 1992. **51**((suppl.)): p. A33 only.
273. Hosler, B.A., et al., *Linkage of familial amyotrophic lateral sclerosis with frontotemporal dementia to chromosome 9q21-q22.* JAMA: The Journal of the American Medical Association, 2000. **284**(13): p. 1664.

274. Hadano, S., et al., *A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2.* Nature Genetics, 2001. **29**(2): p. 166-173.
275. Yang, Y., et al., *The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis.* Nature Genetics, 2001. **29**(2): p. 160-165.
276. Eymard-Pierre, E., et al., *Infantile-onset ascending hereditary spastic paralysis is associated with mutations in the alsin gene.* The American Journal of Human Genetics, 2002. **71**(3): p. 518-527.
277. Al-Chalabi, A., et al., *Variants in the ALS2 gene are not associated with sporadic amyotrophic lateral sclerosis.* Neurogenetics, 2003. **4**(4): p. 221-222.
278. Devon, R., et al., *The first nonsense mutation in alsin results in a homogeneous phenotype of infantileâ€• onset ascending spastic paralysis with bulbar involvement in two siblings.* Clinical Genetics, 2003. **64**(3): p. 210-215.
279. Gros-Louis, F., et al., *An ALS2 gene mutation causes hereditary spastic paraplegia in a Pakistani kindred.* Annals of Neurology, 2003. **53**(1): p. 144-145.
280. Hand, C.K., et al., *Mutation screening of the ALS2 gene in sporadic and familial amyotrophic lateral sclerosis.* Archives of Neurology, 2003. **60**(12): p. 1768.
281. Nagano, I., et al., *Single-nucleotide polymorphisms in uncoding regions of ALS2 gene of Japanese patients with autosomal-recessive amyotrophic lateral sclerosis.* Neurological Research, 2003. **25**(5): p. 505-509.
282. Kress, J.A., et al., *Novel mutation in the ALS2 gene in juvenile amyotrophic lateral sclerosis.* Annals of Neurology, 2005. **58**(5): p. 800-803.
283. Panzeri, C., et al., *The first ALS2 missense mutation associated with JPLS reveals new aspects of alsin biological function.* Brain, 2006. **129**(7): p. 1710.
284. Eymard-Pierre, E., et al., *Novel missense mutation in ALS2 gene results in infantile ascending hereditary spastic paralysis.* Annals of Neurology, 2006. **59**(6): p. 976-980.
285. Verschuuren-Bemelmans, C.C., et al., *Novel homozygous ALS2 nonsense mutation (p. Gln715X) in sibs with infantile-onset ascending spastic paralysis: the first cases from northwestern Europe.* European Journal of Human Genetics, 2008. **16**(11): p. 1407-1411.
286. Shirakawa, K., et al., *Novel compound heterozygous als2 mutations cause juvenile amyotrophic lateral sclerosis in Japan.* Neurology, 2009. **73**(24): p. 2124-2126.
287. Mintchev, N., et al., *A novel ALS2 splice-site mutation in a Cypriot juvenile-onset primary lateral sclerosis family.* Neurology, 2009. **72**(1): p. 28-32.
288. Hirano, M., et al., *Senataxin mutations and amyotrophic lateral sclerosis.* Amyotrophic Lateral Sclerosis, 2011(0): p. 1-5.
289. Chance, P., et al., *Linkage of the gene for an autosomal dominant form of juvenile amyotrophic lateral sclerosis to chromosome 9q34.* The American Journal of Human Genetics, 1998. **62**(3): p. 633-640.
290. Chen, Y., et al., *DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4).* The American Journal of Human Genetics, 2004. **74**(6): p. 1128-1135.
291. Zhao, Z., et al., *A novel mutation in the senataxin gene identified in a Chinese patient with sporadic amyotrophic lateral sclerosis.* Amyotrophic Lateral Sclerosis, 2009. **10**(2): p. 118-122.
292. Orlacchio, A., et al., *SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis.* Brain, 2010. **133**(2): p. 591.
293. Daoud, H., et al., *Exome sequencing reveals< i> SPG11</i> mutations causing juvenile ALS.* Neurobiology of aging, 2012. **33**(4): p. 839. e5-839. e9.
294. Abalkhail, H., et al., *A new familial amyotrophic lateral sclerosis locus on chromosome 16q12. 1-16q12. 2.* The American Journal of Human Genetics, 2003. **73**(2): p. 383-389.
295. Ruddy, D.M., et al., *Two families with familial amyotrophic lateral sclerosis are linked to a novel locus on chromosome 16q.* The American Journal of Human Genetics, 2003. **73**(2): p. 390-396.
296. Sapp, P.C., et al., *Identification of two novel loci for dominantly inherited familial amyotrophic lateral sclerosis.* The American Journal of Human Genetics, 2003. **73**(2): p. 397-403.
297. Belzil, V., et al., *Mutations in FUS cause FALS and SALS in French and French Canadian populations.* Neurology, 2009. **73**(15): p. 1176-1179.
298. Chio, A., et al., *Two Italian kindreds with familial amyotrophic lateral sclerosis due to FUS mutation.* Neurobiology of Aging, 2009. **30**(8): p. 1272-1275.
299. Corrado, L., et al., *Mutations of FUS gene in sporadic amyotrophic lateral sclerosis.* Journal of Medical Genetics, 2009. **47**(3): p. 190.
300. Damme, P.V., et al., *The occurrence of mutations in FUS in a Belgian cohort of patients with familial ALS.* European Journal of Neurology, 2010. **17**(5): p. 754-756.
301. Kwiatkowski Jr, T., et al., *Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis.* Science, 2009. **323**(5918): p. 1205.
302. Ticozzi, N., et al., *Analysis of FUS gene mutation in familial amyotrophic lateral sclerosis within an Italian cohort.* Neurology, 2009. **73**(15): p. 1180.

303. Vance, C., et al., *Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6.* Science, 2009. **323**(5918): p. 1208.
304. Baumer, D., et al., *Juvenile ALS with basophilic inclusions is a FUS proteinopathy with FUS mutations (e-Pub ahead of print).* Neurology, 2010. **75**(7): p. 611.
305. Belzil, V.V., et al., *Identification of a FUS splicing mutation in a large family with amyotrophic lateral sclerosis.* Journal of Human Genetics, 2010. **56**(3): p. 247-249.
306. Blair, I.P., et al., *FUS mutations in amyotrophic lateral sclerosis: clinical, pathological, neurophysiological and genetic analysis.* Journal of Neurology, Neurosurgery & Psychiatry, 2010. **81**(6): p. 639-645.
307. Bosco, D.A., et al., *Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules.* Human molecular genetics, 2010. **19**(21): p. 4160-4175.
308. Broustal, O., et al., *FUS mutations in frontotemporal lobar degeneration with amyotrophic lateral sclerosis.* Journal of Alzheimer's Disease, 2010. **22**(3): p. 765-769.
309. Groen, E.J., et al., *FUS mutations in familial amyotrophic lateral sclerosis in the Netherlands.* Archives of neurology, 2010. **67**(2): p. 224.
310. DeJesus-Hernandez, M., et al., *De novo truncating FUS gene mutation as a cause of sporadic amyotrophic lateral sclerosis.* Human Mutation, 2010. **31**(5): p. E1377-E1389.
311. Huang, E.J., et al., *Extensive FUS-Immunoreactive Pathology in Juvenile Amyotrophic Lateral Sclerosis with Basophilic Inclusions.* Brain Pathology, 2010. **20**(6): p. 1069-1076.
312. Rademakers, R., et al., *Fus gene mutations in familial and sporadic amyotrophic lateral sclerosis.* Muscle and Nerve, 2010. **42**(2): p. 170-176.
313. Suzuki, N., et al., *FALS with FUS mutation in Japan, with early onset, rapid progress and basophilic inclusion.* Journal of Human Genetics, 2010. **55**(4): p. 252-254.
314. Van Langenhove, T., et al., *Genetic contribution of FUS to frontotemporal lobar degeneration.* Neurology, 2010. **74**(5): p. 366-371.
315. Yamamoto-Watanabe, Y., et al., *A Japanese ALS6 family with mutation R521C in the FUS/TLS gene: A clinical, pathological and genetic report.* Journal of the Neurological Sciences, 2010. **296**(1-2): p. 59-63.
316. Yan, J., et al., *Frameshift and novel mutations in FUS in familial amyotrophic lateral sclerosis and ALS/dementia.* Neurology, 2010. **75**(9): p. 807.
317. Belzil, V.V., et al., *Identification of novel FUS mutations in sporadic cases of amyotrophic lateral sclerosis.* Amyotrophic Lateral Sclerosis, 2011(0): p. 1-5.
318. Chio, A., et al., *A de novo missense mutation of the FUS gene in a" true" sporadic ALS case.* Neurobiology of Aging, 2011. **32**(3): p. 553. e23.
319. Drepper, C., et al., *C-terminal FUS/TLS mutations in familial and sporadic ALS in Germany.* Neurobiology of Aging, 2011. **32**(3): p. 548. e1-548. e4.
320. Lai, S.L., et al., *FUS mutations in sporadic amyotrophic lateral sclerosis.* Neurobiology of Aging, 2010. **32**(3): p. 550. e1-550. e4.
321. Robertson, J., et al., *A novel double mutation in FUS gene causing sporadic ALS.* Neurobiology of Aging, 2011. **32**(3): p. 553. e27-553. e30.
322. Syriani, E., M. Morales, and J. Gamez, *FUS/TLS gene mutations are the second most frequent cause of familial ALS in the Spanish population.* Amyotrophic Lateral Sclerosis, 2011. **12**(2): p. 118-123.
323. Belzil, V.V., et al., *Novel FUS Deletion in a Patient With Juvenile Amyotrophic Lateral Sclerosis.* Archives of Neurology, 2012: p. archneurol. 2011.2499 v1.
324. Conte, A., et al., *P525L< i> FUS</i> mutation is consistently associated with a severe form of juvenile Amyotrophic Lateral Sclerosis.* Neuromuscular Disorders, 2012. **22**(1): p. 73-75.
325. Mochizuki, Y., et al., *Familial ALS with< i> FUS</i> P525L mutation: two Japanese sisters with multiple systems involvement.* Journal of the neurological sciences, 2012.
326. Sproviero, W., et al., *< i> FUS</i> mutations in sporadic amyotrophic lateral sclerosis: Clinical and genetic analysis.* Neurobiology of aging, 2012. **33**(4): p. 837. e1-837. e5.
327. Zou, Z.Y., et al., *Screening of the FUS gene in familial and sporadic amyotrophic lateral sclerosis patients of Chinese origin.* European Journal of Neurology, 2012. **19**(7): p. 977-983.
328. van Blitterswijk, M., et al., *Evidence for an oligogenic basis of amyotrophic lateral sclerosis.* Human molecular genetics, 2012. **21**(17): p. 3776-3784.
329. Nishimura, A.L., et al., *A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis.* The American Journal of Human Genetics, 2004. **75**(5): p. 822-831.
330. Landers, J., et al., *New VAPB deletion variant and exclusion of VAPB mutations in familial ALS.* Neurology, 2008. **70**(14): p. 1179.
331. Nishimura, A.L., A. Al-Chalabi, and M. Zatz, *A common founder for amyotrophic lateral sclerosis type 8 (ALS8) in the Brazilian population.* Human Genetics, 2005. **118**(3): p. 499-500.

332. Marques, V.D., et al., *Expanding the phenotypes of the Pro56Ser VAPB mutation: proximal SMA with dysautonomia.* Muscle and Nerve, 2006. **34**(6): p. 731-739.
333. Conforti, F.L., et al., *Sporadic ALS is not associated with VAPB gene mutations in Southern Italy.* Journal of Negative Results in BioMedicine, 2006. **5**(1): p. 7.
334. Funke, A., et al., *The p. P56S mutation in the VAPB gene is not due to a single founder: the first European case.* Clinical Genetics, 2010. **77**(3): p. 302-303.
335. Hayward, C., et al., *Molecular genetic analysis of the APEX nuclease gene in amyotrophic lateral sclerosis.* Neurology, 1999. **Jun 10;52(9)**.
336. Greenway, M., et al., *A novel candidate region for ALS on chromosome 14q11. 2.* Neurology, 2004. **63**(10): p. 1936-1938.
337. Skvortsova, V., et al., *Analysis of heavy neurofilament subunit gene polymorphism in Russian patients with sporadic motor neuron disease (MND).* European Journal of Human Genetics, 2004. **12**(3): p. 241-244.
338. Greenway, M.J., et al., *ANG mutations segregate with familial and'sporadic'amyotrophic lateral sclerosis.* Nature Genetics, 2006. **38**(4): p. 411-413.
339. Crabtree, B., et al., *Characterization of human angiogenin variants implicated in amyotrophic lateral sclerosis.* Biochemistry, 2007. **46**(42): p. 11810-11818.
340. Corrado, L., et al., *Variations in the coding and regulatory sequences of the angiogenin (ANG) gene are not associated to ALS (amyotrophic lateral sclerosis) in the Italian population.* Journal of the Neurological Sciences, 2007. **258**(1-2): p. 123-127.
341. Subramanian, V. and Y. Feng, *A new role for angiogenin in neurite growth and pathfinding: implications for amyotrophic lateral sclerosis.* Human Molecular Genetics, 2007. **16**(12): p. 1445.
342. Wu, D., et al., *Angiogenin lossâ€□ofâ€□function mutations in amyotrophic lateral sclerosis.* Annals of Neurology, 2007. **62**(6): p. 609-617.
343. Conforti, F., et al., *A novel Angiogenin gene mutation in a sporadic patient with amyotrophic lateral sclerosis from southern Italy.* Neuromuscular Disorders, 2008. **18**(1): p. 68-70.
344. Del Bo, R., et al., *Absence of angiogenic genes modification in Italian ALS patients.* Neurobiology of Aging, 2008. **29**(2): p. 314-316.
345. Gellera, C., et al., *Identification of new ANG gene mutations in a large cohort of Italian patients with amyotrophic lateral sclerosis.* Neurogenetics, 2008. **9**(1): p. 33-40.
346. Paubel, A., et al., *Mutations of the ANG gene in French patients with sporadic amyotrophic lateral sclerosis.* Archives of Neurology, 2008. **65**(10): p. 1333.
347. Fernandez-Santiago, R., et al., *Identification of novel Angiogenin (ANG) gene missense variants in German patients with amyotrophic lateral sclerosis.* Journal of Neurology, 2009. **256**(8): p. 1337-1342.
348. Luigetti, M., et al., *SOD1 G93D sporadic amyotrophic lateral sclerosis (SALS) patient with rapid progression and concomitant novel ANG variant.* Neurobiology of Aging, 2011.
349. Kirby, J., et al., *Lack of unique neuropathology in amyotrophic lateral sclerosis associated with p. K54E angiogenin (ANG) mutation.* Neuropathology and applied neurobiology, 2012.
350. DW, D., J. KA, and A.-O. C, *TDP-43 in differential diagnosis of motor neuron disorders.* Acta Neuropathologica, 2007(114): p. 71-79.
351. Daoud, H., et al., *Contribution of TARDBP mutations to sporadic amyotrophic lateral sclerosis.* Journal of Medical Genetics, 2009. **46**(2): p. 112-114.
352. Gitcho, M.A., et al., *TDP-43 A315T mutation in familial motor neuron disease.* Annals of Neurology, 2008. **63**(4): p. 535-538.
353. Kabashi, E., et al., *TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis.* Nature Genetics, 2008. **40**(5): p. 572-574.
354. Kuhnlein, P., et al., *Two German kindreds with familial amyotrophic lateral sclerosis due to TARDBP mutations.* Archives of Neurology, 2008. **65**(9): p. 1185.
355. Rutherford, N., et al., *Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis.* PLoS Genet, 2008. **4**(9): p. e1000193.
356. Sreedharan, J., et al., *TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis.* Science, 2008. **319**(5870): p. 1668.
357. Van Deerlin, V.M., et al., *TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis.* The Lancet Neurology, 2008. **7**(5): p. 409-416.
358. Winton, M.J., et al., *A90V TDP-43 variant results in the aberrant localization of TDP-43 in vitro.* FEBS Letters, 2008. **582**(15): p. 2252-2256.
359. Bäumer, D., N. Parkinson, and K. Talbot, *TARDBP in amyotrophic lateral sclerosis: identification of a novel variant but absence of copy number variation.* Journal of Neurology, Neurosurgery and Psychiatry, 2009. **80**(11): p. 1283.
360. Corrado, L., et al., *High frequency of TARDBP gene mutations in Italian patients with amyotrophic lateral sclerosis.* Human Mutation, 2009. **30**(4): p. 688-694.

361. Del Bo, R., et al., *TARDBP (TDP-43) sequence analysis in patients with familial and sporadic ALS: identification of two novel mutations.* European Journal of Neurology, 2009. **16**(6): p. 727-732.
362. Kamada, M., et al., *Screening for TARDBP mutations in Japanese familial amyotrophic lateral sclerosis.* Journal of the Neurological Sciences, 2009. **284**(1-2): p. 69-71.
363. Kirby, J., et al., *Broad clinical phenotypes associated with TAR-DNA binding protein (TARDBP) mutations in amyotrophic lateral sclerosis.* Neurogenetics, 2010. **11**(2): p. 217-225.
364. Lemmens, R., et al., *TDP-43 M311V mutation in familial amyotrophic lateral sclerosis.* Journal of Neurology, Neurosurgery and Psychiatry, 2009. **80**(3): p. 354.
365. Ticozzi, N., et al., *Mutational analysis of TARDBP in neurodegenerative diseases.* Neurobiology of Aging, 2009.
366. Williams, K., et al., *A novel TARDBP mutation in an Australian amyotrophic lateral sclerosis kindred.* Journal of Neurology, Neurosurgery and Psychiatry, 2009. **80**(11): p. 1286.
367. Iida, A., et al., *Large-scale screening of TARDBP mutation in amyotrophic lateral sclerosis in Japanese.* Neurobiology of Aging, 2010.
368. Praline, J., et al., *Co-occurrence of progressive anarthria with an S393L TARDBP mutation and ALS within a family.* Amyotrophic Lateral Sclerosis, 2012. **13**(1): p. 155-157.
369. Xiong, H.L., et al., *Association between novel TARDBP mutations and Chinese patients with amyotrophic lateral sclerosis.* BMC Medical Genetics, 2010. **11**(1): p. 8.
370. Yokoseki, A., et al., *TDP-43 mutation in familial amyotrophic lateral sclerosis.* Annals of Neurology, 2008. **63**(4): p. 538-542.
371. Borghero, G., et al., *A patient carrying a homozygous p. A382T< i> TARDBP</i> missense mutation shows a syndrome including ALS, extrapyramidal symptoms, and FTD.* Neurobiology of aging, 2011. **32**(12): p. 2327. e1-2327. e5.
372. Chiò, A., et al., *Large proportion of amyotrophic lateral sclerosis cases in Sardinia due to a single founder mutation of the TARDBP gene.* Archives of neurology, 2011: p. archneurol. 2010.352 v1.
373. Fujita, Y., et al., *Different clinical and neuropathologic phenotypes of familial ALS with A315E TARDBP mutation.* Neurology, 2011. **77**(15): p. 1427-1431.
374. Ticozzi, N., et al., *Mutational analysis of< i> TARDBP</i> in neurodegenerative diseases.* Neurobiology of aging, 2011. **32**(11): p. 2096-2099.
375. De Marco, G., et al., *Cytoplasmic accumulation of TDP-43 in circulating lymphomonocytes of ALS patients with and without TARDBP mutations.* Acta Neuropathologica, 2011. **121**(5): p. 611-622.
376. Mosca, L., et al., *Wide phenotypic spectrum of the< i> TARDBP</i> gene: homozygosity of A382T mutation in a patient presenting with amyotrophic lateral sclerosis, Parkinson's disease, and frontotemporal lobar degeneration, and in neurologically healthy subject.* Neurobiology of Aging, 2012. **33**(8): p. 1846. e1-1846. e4.
377. Piaceri, I., et al., *Clinical heterogeneity in Italian patients with amyotrophic lateral sclerosis.* Clinical genetics, 2012. **82**(1): p. 83-87.
378. Zou, Z.-Y., et al., *Screening of the< i> TARDBP</i> gene in familial and sporadic amyotrophic lateral sclerosis patients of Chinese origin.* Neurobiology of Aging, 2012. **33**(9): p. 2229. e11-2229. e18.
379. Chow, C.Y., et al., *Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS.* The American Journal of Human Genetics, 2009. **84**(1): p. 85-88.
380. Maruyama, H., et al., *Mutations of optineurin in amyotrophic lateral sclerosis.* Nature, 2010. **465**(7295): p. 223-226.
381. Belzil, V.V., et al., *Analysis of OPTN as a causative gene for amyotrophic lateral sclerosis.* Neurobiology of Aging, 2010. **32**(3): p. 555. e13-555. e14.
382. Del Bo, R., et al., *Novel optineurin mutations in patients with familial and sporadic amyotrophic lateral sclerosis.* Journal of Neurology, Neurosurgery and Psychiatry, 2011. **82**(11): p. 1239-1243.
383. Iida, A., et al., *Optineurin mutations in Japanese amyotrophic lateral sclerosis.* Journal of Neurology, Neurosurgery and Psychiatry, 2011.
384. Millecamps, S., et al., *Screening of OPTN in French familial amyotrophic lateral sclerosis.* Neurobiology of Aging, 2011. **32**(3): p. 557. e11-557. e13.
385. van Blitterswijk, M., et al., *Novel optineurin mutations in sporadic amyotrophic lateral sclerosis patients.* Neurobiology of Aging, 2011.
386. Tümer, Z., et al., *A novel heterozygous nonsense mutation of the OPTN gene segregating in a Danish family with ALS.* Neurobiology of Aging, 2011.
387. Elden, A.C., et al., *Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS.* Nature, 2010. **466**(7310): p. 1069-1075.
388. Lee, T., et al., *Ataxin-2 intermediate-length polyglutamine expansions in European ALS patients.* Human Molecular Genetics, 2011. **20**(9): p. 1697.
389. Van Damme, P., et al., *Expanded ATXN2 CAG repeat size in ALS identifies genetic overlap between ALS and SCA2.* Neurology, 2011. **76**(24): p. 2066.

390. Koppers, M., et al., *VCP mutations in familial and sporadic amyotrophic lateral sclerosis.* Neurobiology of aging, 2012. **33**(4): p. 837. e7-837. e13.
391. Miller, J.W., et al., *Mutation analysis of< i> VCP</i> in British familial and sporadic amyotrophic lateral sclerosis patients.* Neurobiology of aging, 2012.
392. Ostojic, J., et al., *No evidence of linkage to chromosome 9q21–22 in a Swedish family with frontotemporal dementia and amyotrophic lateral sclerosis.* Neuroscience letters, 2003. **340**(3): p. 245-247.
393. Morita, M., et al., *A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia.* Neurology, 2006. **66**(6): p. 839-844.
394. Vance, C., et al., *Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13. 2–21.3.* Brain, 2006. **129**(4): p. 868-876.
395. Valdmanis, P.N., et al., *Three families with amyotrophic lateral sclerosis and frontotemporal dementia with evidence of linkage to chromosome 9p.* Archives of Neurology, 2007. **64**(2): p. 240.
396. Luty, A.A., et al., *Pedigree with frontotemporal lobar degeneration–motor neuron disease and Tar DNA binding protein-43 positive neuropathology: genetic linkage to chromosome 9.* BMC neurology, 2008. **8**(1): p. 32.
397. Le Ber, I., et al., *Chromosome 9p-linked families with frontotemporal dementia associated with motor neuron disease.* Neurology, 2009. **72**(19): p. 1669-1676.
398. van Es, M.A., et al., *Genome-wide association study identifies 19p13. 3 (UNC13A) and 9p21. 2 as susceptibility loci for sporadic amyotrophic lateral sclerosis.* Nature genetics, 2009. **41**(10): p. 1083-1087.
399. Boxer, A.L., et al., *Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family.* Journal of Neurology, Neurosurgery & Psychiatry, 2011. **82**(2): p. 196-203.
400. Gijselinck, I., et al., *A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study.* The Lancet Neurology, 2012. **11**(1): p. 54-65.
401. Laaksovirta, H., et al., *Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study.* The Lancet Neurology, 2010. **9**(10): p. 978-985.
402. Pearson, J.P., et al., *Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p.* Journal of neurology, 2011. **258**(4): p. 647-655.
403. Shatunov, A., et al., *Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study.* The Lancet Neurology, 2010. **9**(10): p. 986-994.
404. DeJesus-Hernandez, M., et al., *Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS.* Neuron, 2011.
405. Renton, A.E., et al., *A Hexanucleotide Repeat Expansion in< i> C9ORF72</i> Is the Cause of Chromosome 9p21-Linked ALS-FTD.* Neuron, 2011. **72**(2): p. 257-268.
406. Iida, A., et al., *Replication analysis of SNPs on 9p21. 2 and 19p13. 3 with amyotrophic lateral sclerosis in East Asians.* Neurobiology of aging, 2011. **32**(4): p. 757. e13-757. e14.
407. Andersen, P.M., *Mutation in< i> C9orf72</i> changes the boundaries of ALS and FTD.* The Lancet Neurology, 2012. **11**(3): p. 205-207.
408. Bigio, E.H., *C9ORF72, the new gene on the block, causes C9FTD/ALS: new insights provided by neuropathology.* Acta neuropathologica, 2011. **122**(6): p. 653-655.
409. Bigio, E.H., *Motor neuron disease: The C9orf72 hexanucleotide repeat expansion in FTD and ALS.* Nature Reviews Neurology, 2012. **8**(5): p. 249-250.
410. Orr, H.T., *FTD and ALS: genetic ties that bind.* Neuron, 2011. **72**(2): p. 189-190.
411. Wood, H., *Amyotrophic lateral sclerosis: A hexanucleotide repeat expansion in C9ORF72 links amyotrophic lateral sclerosis and frontotemporal dementia.* Nature Reviews Neurology, 2011. **7**(11): p. 595-595.
412. Murray, M.E., et al., *Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72.* Acta neuropathologica, 2011. **122**(6): p. 673-690.
413. Boeve, B.F., et al., *Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72.* Brain, 2012. **135**(3): p. 765-783.
414. Brettschneider, J., et al., *Pattern of ubiquilin pathology in ALS and FTLD indicates presence of C9ORF72 hexanucleotide expansion.* Acta neuropathologica, 2012. **123**(6): p. 825-839.
415. Byrne, S., et al., *Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a< i> C9orf72</i> repeat expansion: a population-based cohort study.* The Lancet Neurology, 2012. **11**(3): p. 232-240.
416. Coon, E.A., et al., *Clinical and electrophysiologic variability in amyotrophic lateral sclerosis within a kindred harboring the C9ORF72 repeat expansion.* Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2013. **14**(2): p. 132-137.

417. Calvo, A., et al., *Amyotrophic lateral sclerosis/frontotemporal dementia with predominant manifestations of obsessive–compulsive disorder associated to GGGGCC expansion of the c9orf72 gene.* Journal of neurology, 2012. **259**(12): p. 2723-2725.
418. Chester, C., et al., *Rapidly progressive frontotemporal dementia and bulbar amyotrophic lateral sclerosis in Portuguese patients with C9orf72 mutation.* Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 2013. **14**(1): p. 70-72.
419. Chio, A., et al., *ALS/FTD phenotype in two Sardinian families carrying both C9ORF72 and TARDBP mutations.* Journal of Neurology, Neurosurgery and Psychiatry, 2012.
420. Cooper-Knock, J., et al., *Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72.* Brain, 2012. **135**(3): p. 751-764.
421. García-Redondo, A., et al., *Analysis of the C9orf72 gene in patients with amyotrophic lateral sclerosis in Spain and different populations worldwide.* Human mutation, 2013. **34**(1): p. 79-82.
422. Herdewyn, S., et al., *Whole-genome sequencing reveals a coding non-pathogenic variant tagging a non-coding pathogenic hexanucleotide repeat expansion in C9orf72 as cause of amyotrophic lateral sclerosis.* Human molecular genetics, 2012. **21**(11): p. 2412-2419.
423. Irwin, D.J., et al., *Cognitive decline and reduced survival in C9orf72 expansion frontotemporal degeneration and amyotrophic lateral sclerosis.* Journal of Neurology, Neurosurgery & Psychiatry, 2013. **84**(2): p. 163-169.
424. Ishiura, H., et al., *C9ORF72 Repeat Expansion in Amyotrophic Lateral Sclerosis in the Kii Peninsula of JapanC9ORF72 Repeat Expansion in ALS.* Archives of neurology, 2012. **69**(9): p. 1154-1158.
425. Jang, J.-H., et al., *Analysis of the< i> C9orf72</i> hexanucleotide repeat expansion in Korean patients with familial and sporadic amyotrophic lateral sclerosis.* Neurobiology of aging, 2012.
426. Lattante, S., et al., *Contribution of major amyotrophic lateral sclerosis genes to the etiology of sporadic disease.* Neurology, 2012. **79**(1): p. 66-72.
427. Majounie, E., et al., *Frequency of the< i> C9orf72</i> hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study.* The Lancet Neurology, 2012. **11**(4): p. 323-330.
428. Millecamps, S., et al., *Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes.* Journal of medical genetics, 2012. **49**(4): p. 258-263.
429. Mok, K.Y., et al., *High frequency of the expanded< i> C9ORF72</i> hexanucleotide repeat in familial and sporadic Greek ALS patients.* Neurobiology of aging, 2012. **33**(8): p. 1851. e1-1851. e5.
430. Ogaki, K., et al., *Analyses of the< i> MAPT</i>,< i> PGRN</i>, and< i> C9orf72</i> mutations in Japanese patients with FTLD, PSP, and CBS.* Parkinsonism & Related Disorders, 2012.
431. Pamphlett, R., et al., *Transmission of C9orf72 hexanucleotide repeat expansions in sporadic amyotrophic lateral sclerosis: an Australian trio study.* NeuroReport, 2012. **23**(9): p. 556-559.
432. Rademakers, R., *C9orf72 repeat expansions in patients with ALS and FTD.* The Lancet Neurology, 2012. **11**(4): p. 297-298.
433. Ratti, A., et al., *< i> C9ORF72</i> repeat expansion in a large Italian ALS cohort: evidence of a founder effect.* Neurobiology of aging, 2012.
434. Rutherford, N.J., et al., *C9ORF72 hexanucleotide repeat expansions in patients with ALS from the Coriell Cell Repository.* Neurology, 2012. **79**(5): p. 482-483.
435. Sabatelli, M., et al., *< i> C9ORF72</i> hexanucleotide repeat expansions in the Italian sporadic ALS population.* Neurobiology of aging, 2012. **33**(8): p. 1848. e15-1848. e20.
436. Savica, R., et al., *Characterization of a family with c9FTD/ALS associated with the GGGGCC repeat expansion in C9ORF72.* Archives of neurology, 2012. **69**(9): p. 1164.
437. Simón-Sánchez, J., et al., *The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions.* Brain, 2012. **135**(3): p. 723-735.
438. Stewart, H., et al., *Clinical and pathological features of amyotrophic lateral sclerosis caused by mutation in the C9ORF72 gene on chromosome 9p.* Acta neuropathologica, 2012. **123**(3): p. 409-417.
439. Beck, J., et al., *Large< i> C9orf72</i> Hexanucleotide Repeat Expansions Are Seen in Multiple Neurodegenerative Syndromes and Are More Frequent Than Expected in the UK Population.* The American Journal of Human Genetics, 2013.
440. Dobson-Stone, C., et al., *Frontotemporal dementia–amyotrophic lateral sclerosis syndrome locus on chromosome 16p12. 1–q12. 2: genetic, clinical and neuropathological analysis.* Acta neuropathologica, 2013: p. 1-11.
441. Gómez-Tortosa, E., et al., *C9ORF72 hexanucleotide expansions of 20–22 repeats are associated with frontotemporal deterioration.* Neurology, 2013. **80**(4): p. 366-370.
442. Jones, A.R., et al., *Residual association at C9orf72 suggests an alternative amyotrophic lateral sclerosis-causing hexanucleotide repeat.* Neurobiology of aging, 2013.

443. Zou, Z.-Y., et al., *Screening for< i> C9orf72</i> repeat expansions in Chinese amyotrophic lateral sclerosis patients.* Neurobiology of aging, 2012.
444. Mitchell, J., et al., *Familial amyotrophic lateral sclerosis is associated with a mutation in D-amino acid oxidase.* Proceedings of the National Academy of Sciences, 2010. **107**(16): p. 7556.
445. Millecamps, S., et al., *Questioning on the role of D amino acid oxidase in familial amyotrophic lateral sclerosis.* Proceedings of the National Academy of Sciences, 2010. **107**(26): p. E107.
446. Puls, I., et al., *Mutant dynactin in motor neuron disease.* Nature Genetics, 2003. **33**(4): p. 455-456.
447. Münch, C., et al., *Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS.* Neurology, 2004. **63**(4): p. 724-726.
448. Münch, C., et al., *Heterozygous R1101K mutation of the DCTN1 gene in a family with ALS and FTD.* Annals of Neurology, 2005. **58**(5): p. 777-780.
449. Figlewicz, D.A., et al., *Variants of the heavy neurofilament subunit are associated with the development of amyotrophic lateral sclerosis.* Human Molecular Genetics, 1994. **3**(10): p. 1757.
450. Rooke, K., et al., *Analysis of the KSP repeat of the neurofilament heavy subunit in familial amyotrophic lateral sclerosis.* Neurology, 1996. **46**(3): p. 789-790.
451. Vechio, J.D., et al., *Sequence variants in human neurofilament proteins: absence of linkage to familial amyotrophic lateral sclerosis.* Annals of Neurology, 1996. **40**(4): p. 603-610.
452. Tomkins, J., et al., *Novel insertion in the KSP region of the neurofilament heavy gene in amyotrophic lateral sclerosis (ALS).* Neuroreport, 1998. **9**(17): p. 3967.
453. Al-Chalabi, A., et al., *Deletions of the heavy neurofilament subunit tail in amyotrophic lateral sclerosis.* Human Molecular Genetics, 1999. **8**(2): p. 157.
454. Daoud, H., et al., *Resequencing of 29 Candidate Genes in Patients With Familial and Sporadic Amyotrophic Lateral Sclerosis.* Archives of Neurology. **68**(5): p. 587.
455. Landers, J., et al., *Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis.* Proceedings of the National Academy of Sciences, 2009. **106**(22): p. 9004.
456. Traynor, B.J., et al., *Kinesin-associated protein 3 (KIFAP3) has no effect on survival in a population-based cohort of ALS patients.* Proceedings of the National Academy of Sciences, 2010. **107**(27): p. 12335-12338.
457. Berg, K.L., T. Seymour, and R. Goel, *History Of Databases.* International Journal of Management & Information Systems (IJMIS), 2012. **17**(1): p. 29-36.
458. Coronel, C. and S.A. Morris, *Database Systems: Design, Implementation, and Management [With Access Code]*. 2012: CengageBrain. com.
459. Spanos, D.-E., P. Stavrou, and N. Mitrou, *Bringing relational databases into the semantic web: A survey.* Semantic Web, 2012. **3**(2): p. 169-209.
460. Doan, A., A. Halevy, and Z. Ives, *Principles of data integration*. 2012: Access Online via Elsevier.
461. Hartmann, S., M. Kirchberg, and S. Link, *Design by example for SQL table definitions with functional dependencies.* The VLDB Journal—The International Journal on Very Large Data Bases, 2012. **21**(1): p. 121-144.
462. Dictionary.com. [cited 2013 20th November]; Available from: http://dictionary.reference.com/browse/genetics?s=t;

http://dictionary.reference.com/browse/database?s=t.
463. Hamosh, A., et al., *Online Mendelian Inheritance in Man (OMIM) Hum.* Mutat, 2000. **15**: p. 57-61.
464. Cotton, R.G., *The HUGO Mutation Database Initiative.* Science, 1998. **279**: p. 10-11.
465. Scriver, C., P. Nowacki, and H. Lehväslaiho, *Guidelines and recommendations for content, structure, and deployment of mutation databases: II. Journey in progress.* Human Mutation, 2000. **15**(1): p. 13-15.
466. Claustres, M., et al., *Time for a unified system of mutation description and reporting: a review of locus-specific mutation databases.* Genome Research, 2002. **12**(5): p. 680.
467. George, R., et al., *General mutation databases: analysis and review.* Journal of Medical Genetics, 2008. **45**(2): p. 65.
468. Soussi, T., et al., *Locus-specific mutation databases: pitfalls and good practice based on the p53 experience.* Nature Reviews Cancer, 2006. **6**(1): p. 83-90.
469. Röhm, U. and J. Blakeley, *Data management for high-throughput genomics.* arXiv preprint arXiv:0909.1764, 2009.
470. Osier, M., et al., *ALFRED: An allele frequency database for anthropology.* American Journal of Physical Anthropology, 2002. **119**(1): p. 77-83.
471. Rajeevan, H., et al., *ALFRED: An allele frequency database for microevolutionary studies.* Evolutionary Bioinformatics Online, 2005. **1**: p. 1.
472. Rajeevan, H., et al., *ALFRED: the ALelle FREquency Database. Update.* Nucleic Acids Research, 2003. **31**(1): p. 270.

473. Krawczak, M., et al., *Human gene mutation database-a biomedical information and research resource.* Human Mutation, 2000. **15**(1): p. 45-51.
474. Stenson, P.D., et al., *The Human Gene Mutation Database: providing a comprehensive central mutation database for molecular diagnostics and personalised genomics.* Human genomics, 2009. **4**(2): p. 69.
475. Fokkema, I.F., et al., *LOVD v. 2.0: the next generation in gene variant databases.* Human mutation, 2011. **32**(5): p. 557-563.
476. Fredman, D., et al., *HGVbase: a curated resource describing human DNA variation and phenotype relationships.* Nucleic Acids Research, 2004. **32**(suppl 1): p. D516.
477. Thorisson, G., et al., *HGVbaseG2P: a central genetic association database.* Nucleic Acids Research, 2009. **37**(Database issue): p. D797.
478. *A haplotype map of the human genome.* Nature, 2005. **437**(7063): p. 1299-320.
479. Thorisson, G., et al., *The international HapMap project web site.* Genome Research, 2005. **15**(11): p. 1592.
480. Sherry, S., et al., *dbSNP: the NCBI database of genetic variation.* Nucleic Acids Research, 2001. **29**(1): p. 308.
481. Glubb, D.M., et al., *A guide to the current web-based resources in pharmacogenomics*, in *Pharmacogenomics*. 2013, Springer. p. 293-310.
482. Chang, H. and T. Fujita, *PicSNP: a browsable catalog of nonsynonymous single nucleotide polymorphisms in the human genome.* Biochemical and Biophysical Research Communications, 2001. **287**(1): p. 288-291.
483. Amberger, J., C. Bocchini, and A. Hamosh, *A new face and new challenges for Online Mendelian Inheritance in Man (OMIM®).* Human mutation, 2011. **32**(5): p. 564-567.
484. Amberger, J., et al., *McKusick's online Mendelian inheritance in man (OMIM®).* Nucleic acids research, 2009. **37**(suppl 1): p. D793-D796.
485. Yoshida, M., et al., *A mutation database for amyotrophic lateral sclerosis.* Human Mutation, 2010. **31**(9): p. 1003-1010.
486. Lill, C.M., et al., *Keeping up with genetic discoveries in amyotrophic lateral sclerosis: The ALSoD and ALSGene databases.* Amyotrophic Lateral Sclerosis, 2011. **12**(4): p. 238-249.
487. Chakrabarti, S., et al., *Data mining curriculum: A proposal (Version 0.91).* 2004.
488. Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth, *From data mining to knowledge discovery in databases.* AI magazine, 1996. **17**(3): p. 37.
489. SIGKDD, A., *Data mining curriculum*, 2012.
490. Berry, M.W. and J. Kogan, *Text mining: applications and theory*. 2010: Wiley. com.
491. Kołcz, A., *Text Mining for Spam Filtering*, in *Encyclopedia of Machine Learning*. 2010, Springer. p. 972-978.
492. Qian, F., et al. *A case for unsupervised-learning-based spam filtering*. in *ACM SIGMETRICS Performance Evaluation Review*. 2010. ACM.
493. Androutsopoulos, I., et al., *Learning to filter spam e-mail: A comparison of a naive bayesian and a memory-based approach.* arXiv preprint cs/0009009, 2000.
494. Wittel, G.L. and S.F. Wu. *On Attacking Statistical Spam Filters*. in *CEAS*. 2004.
495. Ananiadou, S., *Text mining for biomedicine.* Information Retrieval in Biomedicine: Natural Language Processing for Knowledge Integration, 2009: p. 1-10.
496. Hearst, M.A., et al., *BioText Search Engine: beyond abstract search.* Bioinformatics, 2007. **23**(16): p. 2196-2197.
497. Knowles, D.G. and A. McLysaght, *Recent de novo origin of human protein-coding genes.* Genome research, 2009. **19**(10): p. 1752-1759.
498. Ashburner, M., et al., *Gene Ontology: tool for the unification of biology.* Nature genetics, 2000. **25**(1): p. 25-29.
499. Botstein, D., et al., *Gene Ontology: tool for the unification of biology.* Nat Genet, 2000. **25**(1): p. 25-29.
500. Harris, M., et al., *The Gene Ontology (GO) database and informatics resource.* Nucleic acids research, 2004. **32**(Database issue): p. D258-61.
501. Lenzerini, M. *Data integration: A theoretical perspective*. 2002. ACM.
502. Pioro, E. and H. Mitsumoto, *Animal models of ALS.* Clinical neuroscience (New York, NY), 1995. **3**(6): p. 375.
503. Soussi, T., et al., *Locus-specific mutation databases: pitfalls and good practice based on the p53 experience.* Nat Rev Cancer, 2006. **6**(1): p. 83-90.
504. Frazer, K.A., et al., *A second generation human haplotype map of over 3.1 million SNPs.* Nature, 2007. **449**(7164): p. 851-861.
505. Al-Chalabi, A., *Genome-wide association studies.* Cold Spring Harb Protoc, 2009. **2009**(12): p. pdb top66.

506. Moore, J., F. Asselbergs, and S. Williams, *Bioinformatics challenges for genome-wide association studies.* Bioinformatics, 2010. **26**(4): p. 445.
507. Cantor, R., K. Lange, and J. Sinsheimer, *Prioritizing GWAS results: A review of statistical methods and recommendations for their application.* The American Journal of Human Genetics, 2010. **86**(1): p. 6-22.
508. Beleza-Meireles, A. and A. Al-Chalabi, *Genetic studies of amyotrophic lateral sclerosis: controversies and perspectives.* Amyotrophic Lateral Sclerosis, 2009. **10**(1): p. 1-14.
509. Warde-Farley, D., et al., *The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function.* Nucleic Acids Research, 2010. **38**(Web Server issue): p. W214.
510. *Introduction to Computers*. [cited 2013 30th December]; Available from: http://www.cengagesites.com/academic/assets/sites/5096/dc_chapter1.pdf.
511. Bickmore, T.W. and B.N. Schilit, *Digestor: device-independent access to the World Wide Web.* Computer Networks and ISDN Systems, 1997. **29**(8-13): p. 1075-1082.
512. Cutshall, T., L. Blake, and S. Bandy, *Creating a Mobile Library Website.* 2011.
513. Maurer, M.E., et al. *Mobile or desktop websites?: website usage on multitouch devices*. 2010. ACM.
514. Agar, J., *Constant touch: A global history of the mobile phone.* 2005.
515. Iglesias, E. and W. Meesangnil, *Mobile website development: From site to app.* Bulletin of the American Society for Information Science and Technology, 2011. **38**(1): p. 18-23.
516. Imielinski, T. and H. Korth, *Introduction to mobile computing.* Mobile Computing, 1996: p. 1-43.
517. Lemaire, E.D., Y. Boudrias, and G. Greene, *Low-bandwidth, Internet-based videoconferencing for physical rehabilitation consultations.* Journal of Telemedicine and Telecare, 2001. **7**(2): p. 82-89.
518. Gube, J. *8 Tools For Easily Creating a Mobile Version of Your Website*. 2010 [cited 2012 30th October]; Available from: http://mashable.com/2010/12/16/create-mobile-site-tools/.
519. Bosomworth, D. *Mobile Marketing Statistics 2012*. 2012 9th May 2012 [cited 2012 8th November]; Available from: http://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/.
520. Summerfield, J. *Mobile Website vs. Mobile App (Application): Which is Best for Your Organization?* 2012 [cited 2012 28th August]; Available from: http://www.hswsolutions.com/services/mobile-web-development/mobile-website-vs-apps/.
521. Abel, O., et al., *ALSoD: A user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics.* Human Mutation, 2012.
522. Abel, O., et al., *Credibility Analysis of Putative Disease-Causing Genes Using Bioinformatics.* PloS one, 2013. **8**(6): p. e64899.
523. Abel, O., et al., *Development of a Smartphone App for a Genetics Website: The Amyotrophic Lateral Sclerosis Online Genetics Database (ALSoD).* Journal of Medical Internet Research, 2013. **15**(9).
524. Homer, N., et al., *Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays.* PLoS Genet, 2008. **4**(8): p. e1000167.
525. White, J., et al., *Guidelines for human gene nomenclature (1997). HUGO Nomenclature Committee.* Genomics, 1997. **45**(2): p. 468.
526. Maglott, D., et al., *Entrez Gene: gene-centered information at NCBI.* Nucleic Acids Research. **39**(suppl 1): p. D52.
527. Fujita, P.A., et al., *The UCSC Genome Browser database: update 2011.* Nucleic Acids Research. **39**(suppl 1): p. D876.
528. Rose, P.W., et al., *The RCSB Protein Data Bank: redesigned web site and web services.* Nucleic Acids Research. **39**(suppl 1): p. D392.
529. Amberger, J., C. Bocchini, and A. Hamosh, *A new face and new challenges for Online Mendelian Inheritance in Man (OMIMÂ®).* Human Mutation, 2011.
530. Amberger, J., et al., *McKusick's Online Mendelian Inheritance in Man (OMIM®).* Nucleic Acids Research, 2009. **37**(suppl 1): p. D793.
531. Safran, M., et al., *GeneCards Version 3: the human gene integrator.* Database: the journal of biological databases and curation. **2010**.
532. Gasteiger, E., et al., *Protein identification and analysis tools on the ExPASy server.* The proteomics protocols handbook, 2005: p. 571-607.
533. Kanehisa, M., et al., *KEGG for representation and analysis of molecular networks involving diseases and drugs.* Nucleic Acids Research, 2000. **38**(suppl 1): p. D355.
534. Jain, E., et al., *Infrastructure for the life sciences: design and implementation of the UniProt website.* BMC Bioinformatics, 2009. **10**(1): p. 136.
535. Hoffmann, R. and A. Valencia, *A gene network for navigating the literature.* Nature Genetics, 2004. **36**(7): p. 664-664.

536. Pagon, R.A., et al., *GeneTests GeneClinics: Genetic testing information for a growing audience.* Human Mutation, 2002. **19**(5): p. 501-509.
537. Carbon, S., et al., *AmiGO: online access to ontology and annotation data.* Bioinformatics, 2009. **25**(2): p. 288.
538. Hubbard, T., et al., *Ensembl 2009.* Nucleic Acids Research, 2009. **37**(suppl 1): p. D690.
539. Lill, C., et al., *The ALSGene Database: Systematic Meta-Analyses and Field Synopsis of Genetic Association Studies in ALS*, in *21st International Symposium on ALS/MND*2010. p. Abstract #SW264.
540. Huss III, J.W., et al., *A gene wiki for community annotation of gene function.* PLoS biology, 2008. **6**(7): p. e175.
541. Maret, S., *Wolfram| Alpha Knowledgebase.* The Charleston Advisor. **12**(2): p. 48-50.
542. Hoffmann, R., *A wiki for the life sciences where authorship matters.* Nature Genetics, 2008. **40**(9): p. 1047-1051.
543. Hindorff, L.A., et al., *Potential etiologic and functional implications of genome-wide association loci for human diseases and traits.* Proceedings of the National Academy of Sciences, 2009. **106**(23): p. 9362.
544. Fogh, I., et al., *A genome-wide association meta-analysis identifies a novel locus at 17q11. 2 associated with sporadic amyotrophic lateral sclerosis.* Human molecular genetics, 2013: p. ddt587.
545. Gold, M. ***Creating a Blog in ASP.NET and C#***. 2005.
546. Stenson, P.D., et al., *Human Gene Mutation Database: towards a comprehensive central mutation database.* Journal of Medical Genetics, 2008. **45**(2): p. 124-6.
547. Galperin, M.Y. and X.M. FernÃ¡ndez-SuÃ¡rez, *The 2012 Nucleic Acids Research Database Issue and the online Molecular Biology Database Collection.* Nucleic Acids Research. **40**(D1): p. D1-D8.
548. Thomas, P.D., et al., *PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification.* Nucleic Acids Research, 2003. **31**(1): p. 334.
549. Ng, P.C. and S. Henikoff, *SIFT: Predicting amino acid changes that affect protein function.* Nucleic Acids Research, 2003. **31**(13): p. 3812.
550. Ramensky, V., P. Bork, and S. Sunyaev, *Human non-synonymous SNPs: server and survey.* Nucleic Acids Research, 2002. **30**(17): p. 3894-3900.
551. Waterhouse, A.M., et al., *Jalview Version 2—a multiple sequence alignment editor and analysis workbench.* Bioinformatics, 2009. **25**(9): p. 1189.
552. Barrett, J.C., *Haploview: Visualization and analysis of SNP genotype data.* Cold Spring Harbor Protocols, 2009. **2009**(10): p. pdb. ip71.
553. Barrett, J.C., et al., *Haploview: analysis and visualization of LD and haplotype maps.* Bioinformatics, 2005. **21**(2): p. 263-265.
554. Adzhubei, I., D.M. Jordan, and S.R. Sunyaev, *Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2.* Current Protocols in Human Genetics, 2013: p. 7.20. 1-7.20. 41.
555. Kenna, K.P., et al., *Using Reference Databases of Genetic Variation to Evaluate the Potential Pathogenicity of Candidate Disease Variants.* Human mutation, 2013.
556. Cooper, D.N. and M. Krawczak, *Human gene mutation database.* Human genetics, 1996. **98**(5): p. 629-629.
557. Cornell, M. and A. Devereau. *The First GEN2PHEN Pilot Project: Documenting the bioinformatics requirements of clinical scientists*. in *JOURNAL OF MEDICAL GENETICS*. 2009. BMJ PUBLISHING GROUP BRITISH MED ASSOC HOUSE, TAVISTOCK SQUARE, LONDON WC1H 9JR, ENGLAND.
558. Cornell, M., G. Dobson, and A. Devereau. *Overview of GEN2PHEN software developments for analysis and management of genetic variant data*. in *JOURNAL OF MEDICAL GENETICS*. 2010. BMJ PUBLISHING GROUP BRITISH MED ASSOC HOUSE, TAVISTOCK SQUARE, LONDON WC1H 9JR, ENGLAND.
559. Aymé, S., *Orphanet, an information site on rare diseases].* Soins; la revue de référence infirmière, 2003(672): p. 46.
560. Weinreich, S., et al., *Orphanet: a European database for rare diseases].* Nederlands tijdschrift voor geneeskunde, 2008. **152**(9): p. 518.
561. Wu, C.-C., et al., *GeneNetwork: an interactive tool for reconstruction of genetic networks using microarray data.* Bioinformatics, 2004. **20**(18): p. 3691-3693.
562. Bult, C.J., et al., *The Mouse Genome Database (MGD): mouse biology and model systems.* Nucleic acids research, 2008. **36**(suppl 1): p. D724-D728.
563. Eppig, J.T., et al., *The Mouse Genome Database (MGD): from genes to mice—a community resource for mouse biology.* Nucleic acids research, 2005. **33**(suppl 1): p. D471-D475.
564. Fokkema IF, et al., *LOVD v.2.0: The Next Generation in Gene Variant Databases.* Human Mutation, 2011. **32**(5): p. 557-63.

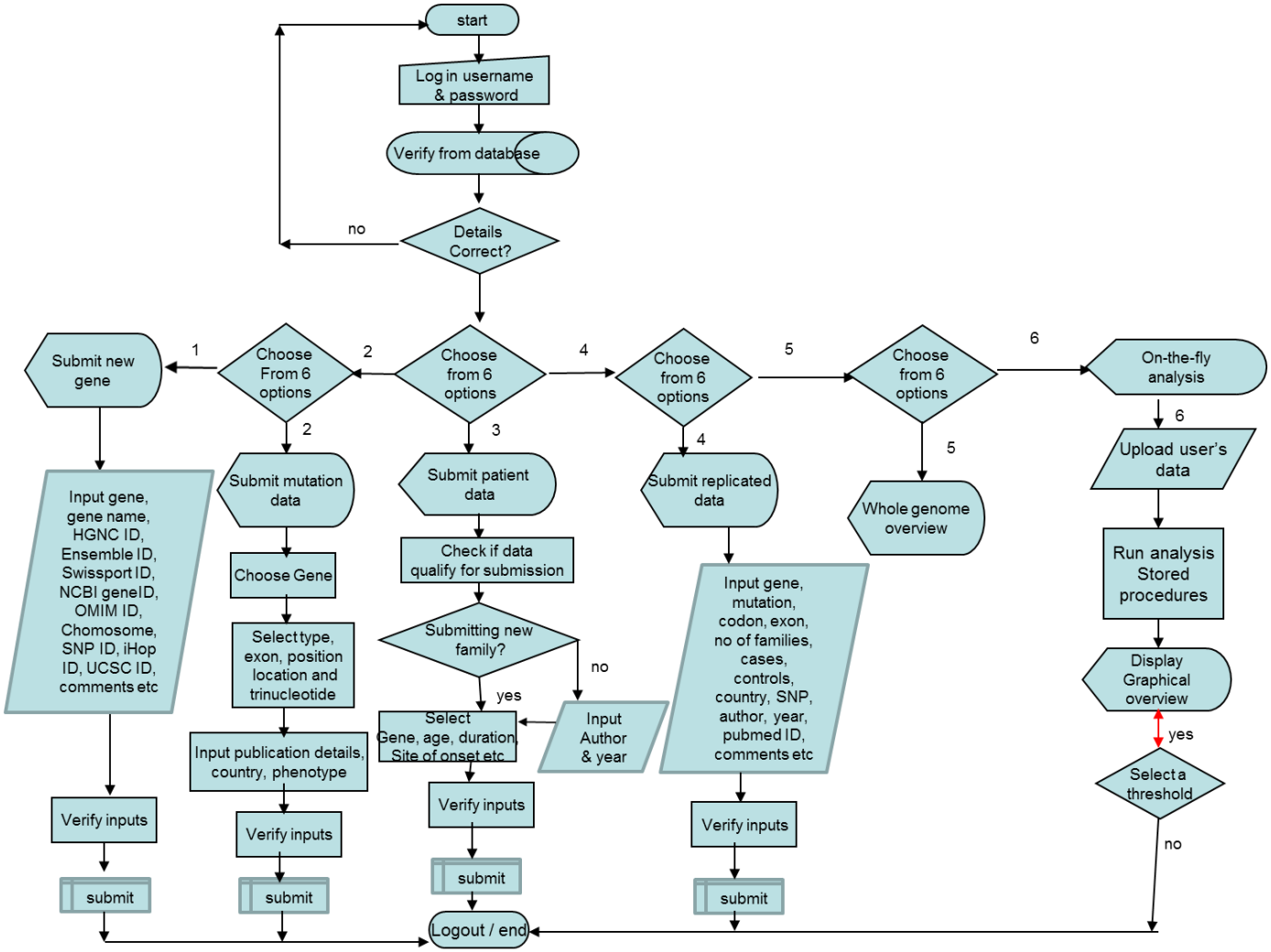
565. Fokkema, I.F., J.T. den Dunnen, and P.E. Taschner, *LOVD: Easy creation of a locus-specific sequence variation database using an "LSDB-in-a-box" approach.* Human mutation, 2005. **26**(2): p. 63-68.
566. Li, K.B., *ClustalW-MPI: ClustalW analysis using distributed and parallel computing.* Bioinformatics, 2003. **19**(12): p. 1585.
567. Radunovic, A. and P. Leigh, *ALSODatabase: Database of SOD1 (and other) gene mutations in ALS on the Internet.* Amyotrophic Lateral Sclerosis, 1999. **1**(1): p. 45-49.
568. Wroe, R., et al., *ALSOD: the amyotrophic lateral sclerosis online database.* Amyotrophic Lateral Sclerosis, 2008. **9**(4): p. 249-250.
569. Sunyaev, S., et al., *Prediction of deleterious human alleles.* Human Molecular Genetics, 2001. **10**(6): p. 591.
570. !!! INVALID CITATION !!!
571. Agency for Toxic Substances and Disease Registry. *National Amyotrophic Lateral Sclerosis (ALS) Registry*. 2011; Available from: https://wwwn.cdc.gov/als/ALSFAQ.aspx#j.
572. David, F. and C. Mallows, *The variance of Spearman's rho in normal samples.* Biometrika, 1961. **48**(1/2): p. 19-28.
573. Ramsey, P.H., *Critical values for Spearman's rank order correlation.* Journal of Educational and Behavioral Statistics, 1989. **14**(3): p. 245-253.
574. Yue, S., P. Pilon, and G. Cavadias, *Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series.* Journal of Hydrology, 2002. **259**(1): p. 254-271.
575. Flow, A. *How to Make a Mobile Website – 6 Easy Tips*

2010 [cited 2012 30th October]; Available from: http://blog.templatemonster.com/2010/05/11/how-make-mobile-website-6-easy-tips/.
576. Georgiou, N. ***ASP.NET IsMobileDevice and how it works***

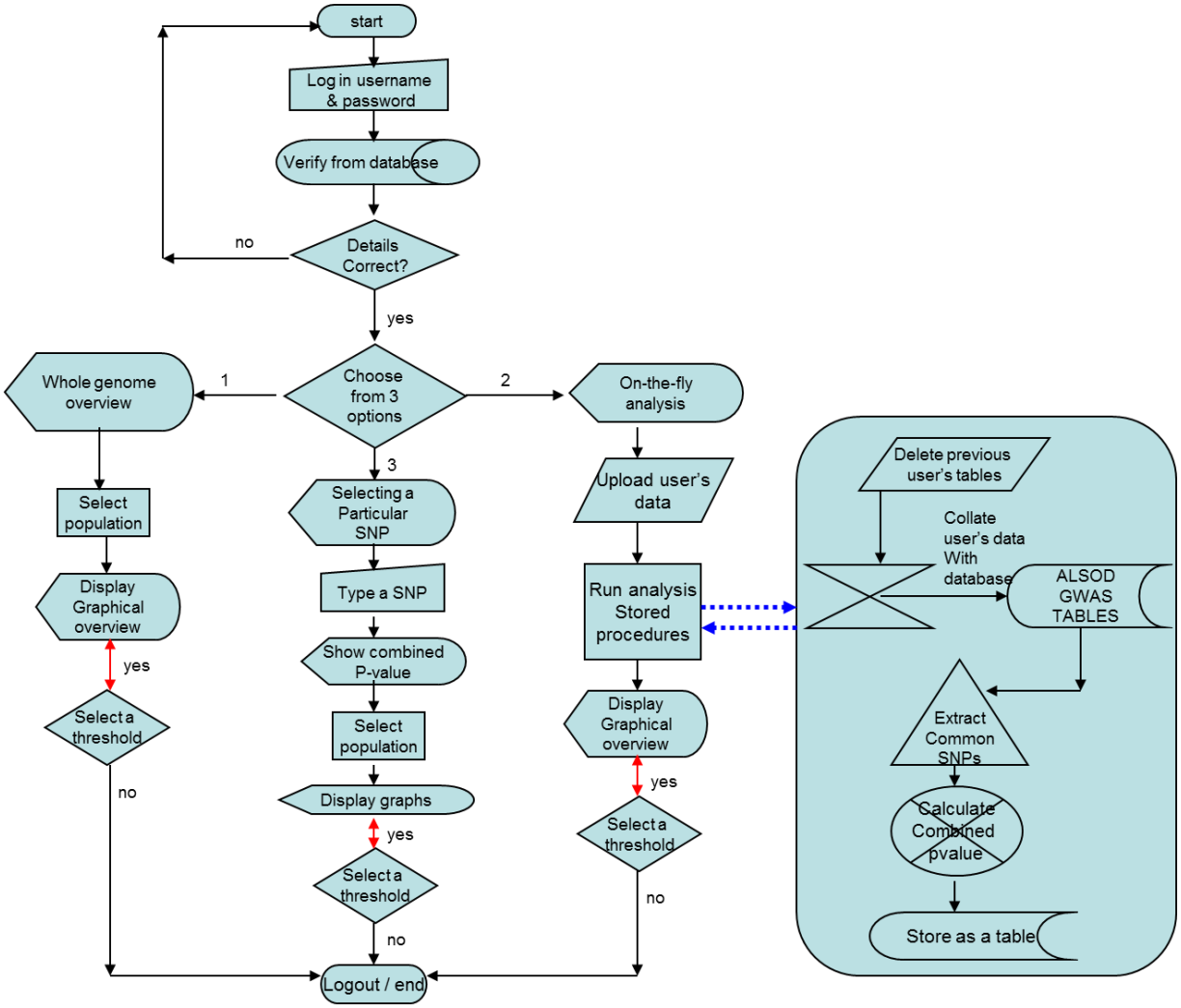
2011 [cited 2012 30th October]; Available from: http://ngeor.net/blog/post/2011/03/12/asp-net-IsMobileDevice-and-how-it-works.aspx.
577. Hamblen M. *Android fastest growing smartphone OS in U.S., survey finds.* 2010 [cited 2012 24th October]; Available from: http://www.computerworld.com/s/article/9199499/Android_fastest_growing_smartphone_OS_in_U.S._survey_finds.
578. ***Building Your First App***. 2012 [cited 2012 1st October 2012]; Available from: http://developer.android.com/training/basics/firstapp/index.html.
579. ***Web Apps***. 2012 [cited 2012 12th October]; Available from: http://developer.android.com/guide/webapps/index.html.
580. Heuwieser, F.M. *How to write your first BlackBerry application – Lesson 1: UI fields*. 2009 [cited 2012 1st October 2012]; Available from: http://www.blackberrycool.com/2009/06/05/how-to-write-your-first-blackberry-application-lesson-1-ui-fields/.
581. Genuitec. *Genuitec Product and Eclipse Package Downloads*. 2012 [cited 2012 15th October]; Available from: http://www.genuitec.com/eclipse_packages.html.
582. *BlackBerry Java 7.1 SDK*. BlackBerry Plug-in for Eclipse 2012 [cited 2012 15th October]; Available from: https://developer.blackberry.com/java/download/eclipse.
583. Greenfield, M. *How to develop Android apps*. 2012 [cited 2012 8th October]; Available from: http://reviews.cnet.co.uk/mobile-apps/how-to-develop-android-apps-50008626/.
584. OBryan, S. *Android - Use websites in your app (WebView Tutorial)* Fun Projects of Various Types 2011 [cited 2011 August 27]; Available from: http://techvalleyprojects.blogspot.co.uk/2011/08/android-use-websites-in-your-app.html.
585. Hickson, I., *A vocabulary and associated APIs for HTML and XHTML.* World Wide Web Consortium, Working Draft WD-html5-20120329, 2012.
586. vandenbossche, J. *HTML5 Application Cache study*. [cited 2012 19th October]; Available from: http://www.winktoolkit.org/blog/235/.
587. Grove, R. *Mobile Browser Cache Limits: Android, iOS, and webOS*. 2010 [cited 2012 19th October]; Available from: http://www.yuiblog.com/blog/2010/06/28/mobile-browser-cache-limits/.
588. Smith, S. *Android SDK: Implement an Options Menu*. 2012 [cited 2012 2nd November]; Available from: http://mobile.tutsplus.com/tutorials/android/android-sdk-implement-an-options-menu/.
589. Developers, G. *Google App Engine*. 2012 [cited 2012 2nd October 2012]; Available from: https://developers.google.com/appengine/docs/whatisgoogleappengine.
590. Plaza, B., *Google Analytics for measuring website performance.* Tourism Management, 2011. **32**(3): p. 477-481.
591. Hindorff, L., et al., *A catalog of published genome-wide association studies.* National Human Genome Research Institute (Available at http://www. genome. gov/gwastudies.).
592. Yu, Y., et al., *HuGE Literature Finder. HuGE Navigator. 2008*.

593. Moher, D., et al., *Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.* PLoS medicine, 2009. **6**(7): p. e1000097.
594. Radunovic, A. and P. Leigh, *Cu/Zn superoxide dismutase gene mutations in amyotrophic lateral sclerosis: correlation between genotype and clinical features.* Journal of Neurology, Neurosurgery and Psychiatry, 1996. **61**(6): p. 565.
595. ALSoD. *ALS Online genetics Database*. Available from: http://alsod.iop.kcl.ac.uk.
596. Ticozzi, N., et al., *Genetics of familial amyotrophic lateral sclerosis.* Archives Italiennes de Biologie, 2011. **149**(1): p. 65-82.
597. Fisher, R., *Combining independent tests of significance.* American Statistician, 1948. **2**(5): p. 30.
598. Singh, G., *Conventional Approach Of Combining P-Values Using Chi-Square Test: An Essence of Meta Analysis.* The Internet Journal of Medical Informatics, 2006. **3**(1).
599. Epstein, M., A. Allen, and G. Satten, *A simple and improved correction for population stratification in case-control studies.* The American Journal of Human Genetics, 2007. **80**(5): p. 921-930.
600. Boucher, B. and S. Jenna, *Genetic interaction networks: better understand to better predict.* Frontiers in genetics, 2013. **4**.
601. Claustres, M., et al., *Time for a unified system of mutation description and reporting: a review of locus-specific mutation databases.* Genome Research, 2002. **12**(5): p. 680-688.
602. Egea, R., et al., *MamPol: a database of nucleotide polymorphism in the Mammalia class.* Nucleic Acids Research, 2006.
603. Najmabadi, H., et al., *The Iranian Human Mutation Gene Bank: a data and sample resource for worldwide collaborative genetics research.* Human Mutation, 2003. **21**(2): p. 146-150.
604. Cotton, R.G., *Progress of the HUGO mutation database initiative: a brief introduction to the human mutation MDI special issue.* Human Mutation, 2000. **15**(1): p. 4-6.
605. Van Deerlin, V.M., et al., *Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions.* Nat Genet, 2010. **42**(3): p. 234-9.
606. Vance, C., et al., *Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3.* Brain, 2006. **129**(Pt 4): p. 868-76.
607. Morita, M., et al., *A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia.* Neurology, 2006. **66**(6): p. 839-44.
608. Shatunov, A., et al., *Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study.* Lancet Neurol. **9**(10): p. 986-94.
609. Laaksovirta, H., et al., *Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study.* Lancet Neurol. **9**(10): p. 978-85.
610. Al-Chalabi, A., D. Aarsland, and A. Rosser, *Journal Watch: Our panel of experts highlight the most important research articles across the spectrum of topics relevant to the field of neurodegenerative disease management.* Neurodegenerative Disease Management, 2012. **2**(1): p. 103-105.
611. Zheng, X., *A High-performance Computing Toolset for Big Data Analysis of Genome-Wide Variants.* 2013.
612. Wicks, P., et al., *Sharing health data for better outcomes on PatientsLikeMe.* Journal of medical Internet research, 2010. **12**(2).
613. Moreira, L.G.A., et al., *Structural and Functional Analysis of Human SOD1 in Amyotrophic Lateral Sclerosis.* PLOS ONE, 2013. **8**(12): p. e81979.
614. Dion, P.A., H. Daoud, and G.A. Rouleau, *Genetics of motor neuron disorders: new insights into pathogenic mechanisms.* Nature Reviews Genetics, 2009. **10**(11): p. 769-782.
615. Lambrechts, D., et al., *Meta-analysis of VEGF variations in ALS: increased susceptibility in male carriers of the-2578AA genotype.* Journal of Medical Genetics, 2008.
616. Bertram, L., et al., *The AlzGene Database. Alzheimer Research Forum.* Avalaible at: http://www.alzgene.org, 2005.
617. Lill, C., et al., *The MSGene Database*, Alzheimer Research Forum.
618. Lill, C., et al., *The PDGene Database*, 2001, Alzheimer Research Forum.
619. Allen, N., et al., *Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database.* Nature Genetics, 2008. **40**(7): p. 827-834.

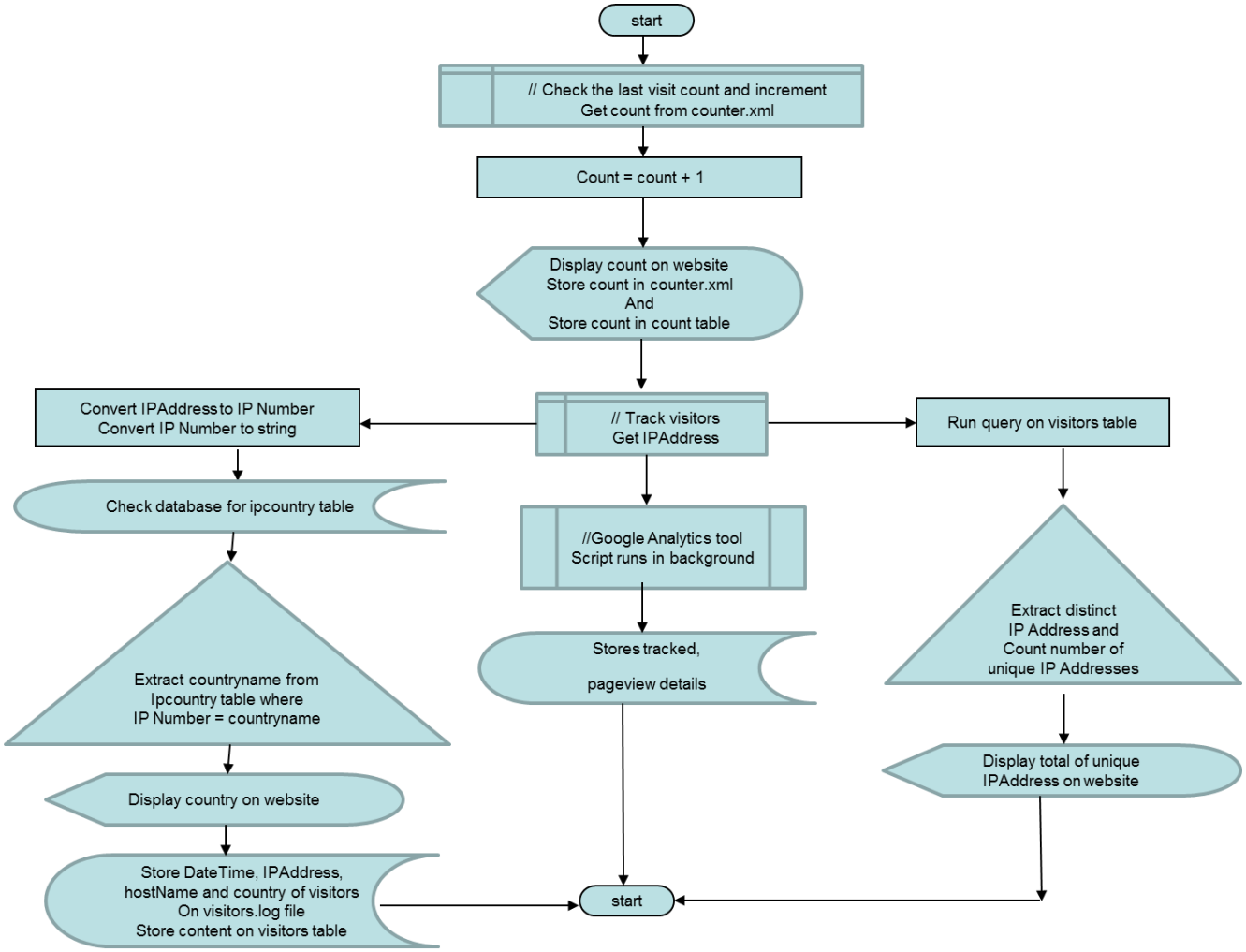
Appendix 1 – Data Submission Flowchart



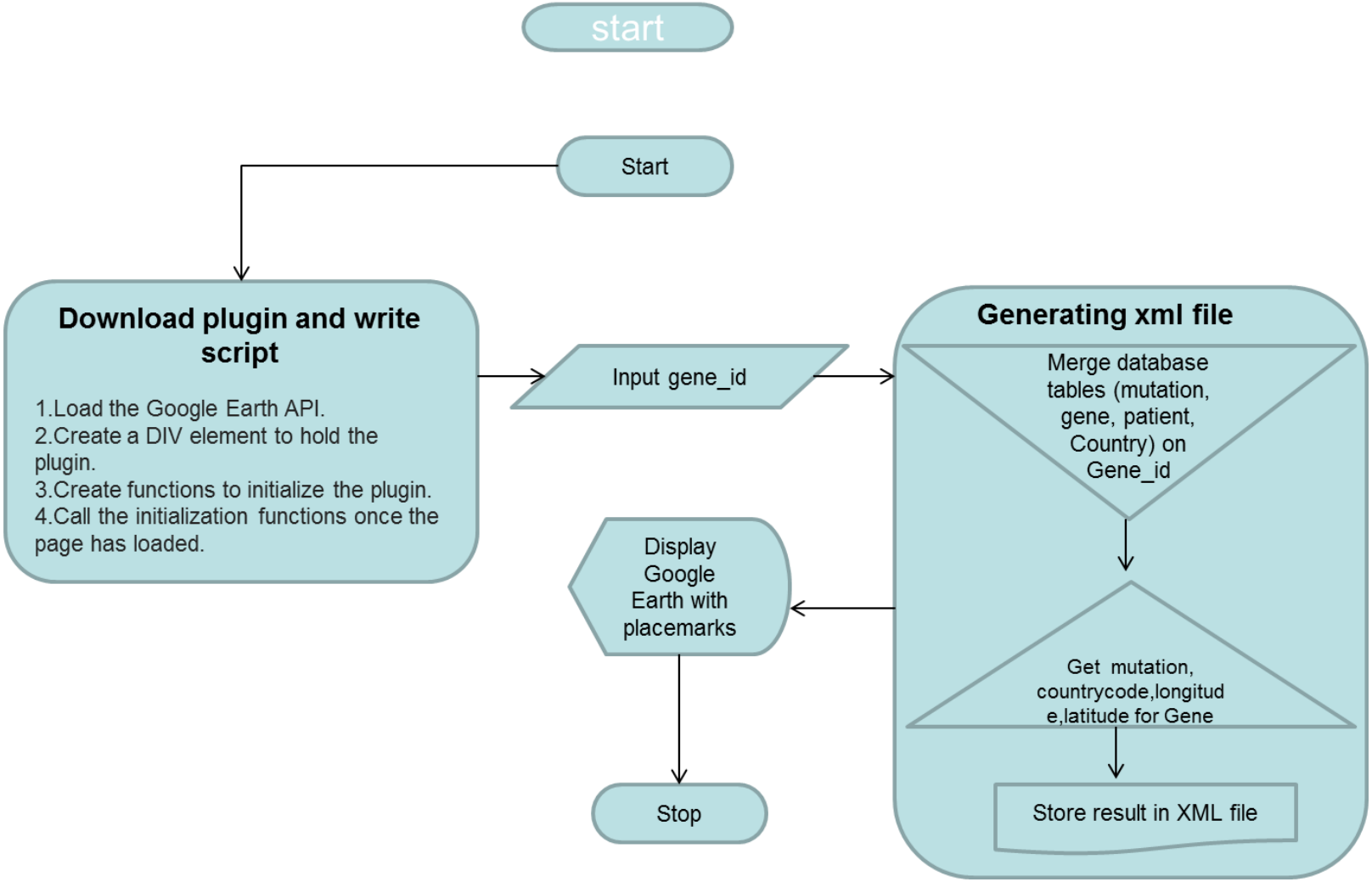
Appendix 2 – Genome Wide Association Study Flowchart



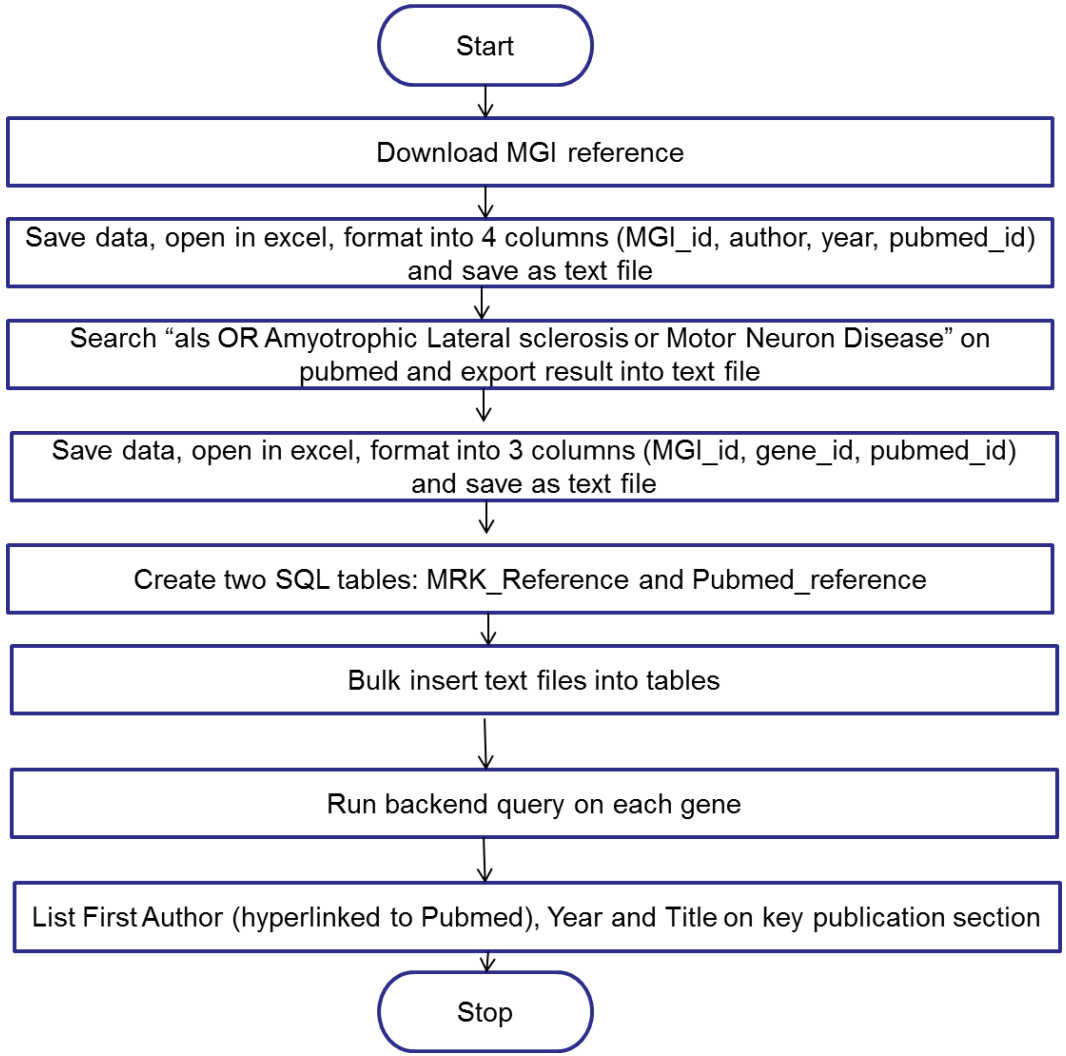
Appendix 3 – Visitors' Statistics



Appendix 4 – Embedding Google Earth Flowchart



Appendix 5 – Flowchart for integrating Animal model into ALSoD



Appendix 6 - Data extraction from publications

Review protocol

Primary database: PubMed (http://www.ncbi.nlm.nih.gov/pubmed/) and Google Scholar (http://scholar.google.co.uk/)

Search terms on PubMed:

(SOD1[Title] OR (superoxide dismutase[Title]) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) = 181 results
ALS2 [Title] OR Alsin[Title] AND (mutation[Title] OR novel[Title]) AND (Amyotrophic Lateral Sclerosis[Title] OR Motor Neuron Disease[Title] OR ALS[Title]) = 6 results
ANG [Title] OR Angiogenin[Title] AND (mutation[Title] OR novel[Title]) AND (Amyotrophic Lateral Sclerosis[Title] OR Motor Neuron Disease[Title] OR ALS[Title]) = 6 results
(FUS [Title] OR (fusion [Title]) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) = 21 results
((TARDBP [Title]) OR (TDP-43 [Title]) OR (fusion [Title])) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) = 24 results
(VAPB [Title]) OR (Vesicle-associated membrane [Title])) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) = 6 results
(NEFH [Title]) OR (neurofilament [Title]) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) = 1 results
(SPG11 [Title] AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) = 0 result
OPTN [Title] OR optineurin [Title] AND (mutation[Title] OR novel[Title]) AND (Amyotrophic Lateral Sclerosis[Title] OR Motor Neuron Disease[Title] OR ALS[Title]) = 11 results

SETX [Title] OR Senataxin [Title] AND (mutation[Title] OR novel[Title]) AND (Amyotrophic Lateral Sclerosis[Title] OR Motor Neuron Disease[Title] OR ALS[Title]) = 2 results

FIG4 [Title] AND (mutation[Title] OR novel[Title]) AND (Amyotrophic Lateral Sclerosis[Title] OR Motor Neuron Disease[Title] OR ALS[Title]) = 0

DCTN1 [Title] OR Dynactin [Title] AND (mutation[Title] OR novel[Title]) AND (Amyotrophic Lateral Sclerosis[Title] OR Motor Neuron Disease[Title] OR ALS[Title]) = 3 results

TAF15 [Title] AND (mutation[Title] OR novel[Title]) AND (Amyotrophic Lateral Sclerosis[Title] OR Motor Neuron Disease[Title] OR ALS[Title]) = 1 results

(VCP [Title] OR (valosin-containing protein [Title]) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) = 4 results

(DAO [Title] OR (D-amino-acid oxidase [Title]) AND (mutation[Title] OR novel [Title])AND ((Amyotrophic Lateral Sclerosis[Title]) OR (Motor Neuron Disease[Title]) OR ALS[Title]) = 1 result

Search terms on Google Scholar:

SOD1 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 2050 results

ALS2 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 546 results

ANG novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 470 results

FUS novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 2520 results

TARDBP TDP-43 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 608 results

VAPB novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 396 results

NEFH novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 79 results

SPG11 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 53 results

OPTN novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 1850 results

SETX novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 1020 results

FIG4 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 930 results

DCTN1 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 261 results
TAF15 novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 68 results
VCP novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 404 results
DAO novel mutations variants ALS "amyotrophic lateral sclerosis" "motor neuron disease" = 52 results

Additional databases:

The Human Gene Mutation Database at the Institute of Medical Genetics in Cardiff on http://www.hgmd.cf.ac.uk/ac/index.php
ALSGene by the Max Planck Institute for Molecular Genetics Berlin, the Alzheimer Research Forum and Prize4Life on http://www.alzgene.org/
ALS mutation database supported by the Ministry of Education Japan on http://reseq.biosciencedbc.jp/resequence/SearchDisease.do?targetId=1
Online Mendelian Inheritance in Man (OMIM) on http://www.ncbi.nlm.nih.gov/omim

Inclusion criteria:

Amyotrophic Lateral Sclerosis
Motor Neuron Disease
Mutations/ variants

Exclusion criteria:

Animal models
Associated with other diseases
Patients already examined in another study

Data abstraction forms

Identification of data abstractor:

Log in with username and password

Add a mutation

Select gene

Select mutation location

Select mutation type

Select sequence position and trinucleotide

Select new trinucleotide after mutation

Select zygosity

Select mutation documentation

Input first author

Input year of publication

Input paper title

Input full paper link

Input Doi key (where available)

Input country(s) where mutation is found

Selected Phenotype by default (Amyotrophic Lateral Sclerosis)

Input dbSNP

Submit a patient

Conduct preliminary screen to check if patient data needs to be submitted to database (like is subject affected?, any mutation found?, any mutations in subject's family members?)

Select or add a new family id (represented by first author and year e.g Smith 2009)

Select gender

Select country of origin (where not available, consider where research was conducted)

Select ethnic origin (where available)

Select if dead or alive

Choose gene

Has patient been screened?

Was mutation found?

Is there a family history?

Select affected or unaffected status.

Select zygosity

Select mutation

Select site of onset and side of the body

Input age of onset

Input disease duration (in months)

Select UMN or LMN or Cognitive signs

Select phenotype (ALS or FTD or ALS-FTD or Unknown)

Submit a gene

Input Gene ID (e.g SOD1)

Input HGNC ID (e.g. 11179)

Input Ensemble ID (e.g. ENSG00000142168)

Input Swissport ID (e.g. P00441)

Input NCBI gene ID (e.g. 6647)

Input NCBI refseq ID (e.g. NM_000454)

Input Structure ID (e.g. uc002ypa.1)

Input OMIM ID (e.g. 147450)

Input Genecards ID (e.g. SOD1)

Input Gene Name (e.g. Cu/Zn superoxide dismutase 1,...)

Input Keywords (e.g. SOD1)

Input Chromosome Name (e.g. 21)

Input Chromosome Position(e.g. q)

Input Chromosome Band (e.g. 22)

Input Chromosome Bp (e.g. 11)

Input Chromosome (e.g. 21q22.11)

Input Protein Name (e.g. superoxide dismutase 1, soluble)

Input Protein Function (e.g.destroys radicals which are....)

Input Phenotype (e.g. defects in sod1 are the cause of ALS...)

Input Gene comments (e.g. gene: sod1. [153 amino acids; 15 kd])

Input Other_names (e.g. ALS1)

Input Accession ID (e.g. AY04978)

Input reason for investigation (e.g. Mutations in SOD1 account for 20% of familial ALS)

Input Result (e.g. 2-7% sporadic cases have mutations)

Input category (e.g OXIDATIVE STRESS)

Input Gene effect (e.g. FALS genes found in SALS)

Input iHop (e.g. 92317)

Input pdb_id (e.g. 2C9V)

Input dbSNP (e.g. rs92317)

Input pubmed_id

Submit a replicated mutation

Input Gene ID

Input Mutation

Input Codon

Input Exon

Input Number of Independent families

How many generations were examined?

How many PATIENTS were affected with the mutation? e.g. 5

How many CONTROLS were affected with the mutation? e.g. 0

How many Sporadic ALS patients were examined? e.g. 55

How many Familial ALS patients were examined? e.g 50

Is this mutation reported as SALS or FALS?

Total number of cases examined e.g 15

Total number of controls examined e.g 10

LOD Score reported e.g 6.66

Is a pedigree or Linkage analysis shown?

Input Country of origin

Input Ethnic Origin :

Input SNP (e.g. rs11010):

Is mutation pathogenic ?

Input first Author

Input Year

Input Pubmed ID (e.g. 2020294)

Characteristics of extracted data:

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SOD1

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=ALS2

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=ANG

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=FUS

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=TARDBP(TDP43)

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=VAPB

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=NEFH

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SPG11

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=OPTN

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=SETX

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=FIG4

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=DCTN1

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=TAF15

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=VCP

http://alsod.iop.kcl.ac.uk/Overview/gene.aspx?gene_id=DAO

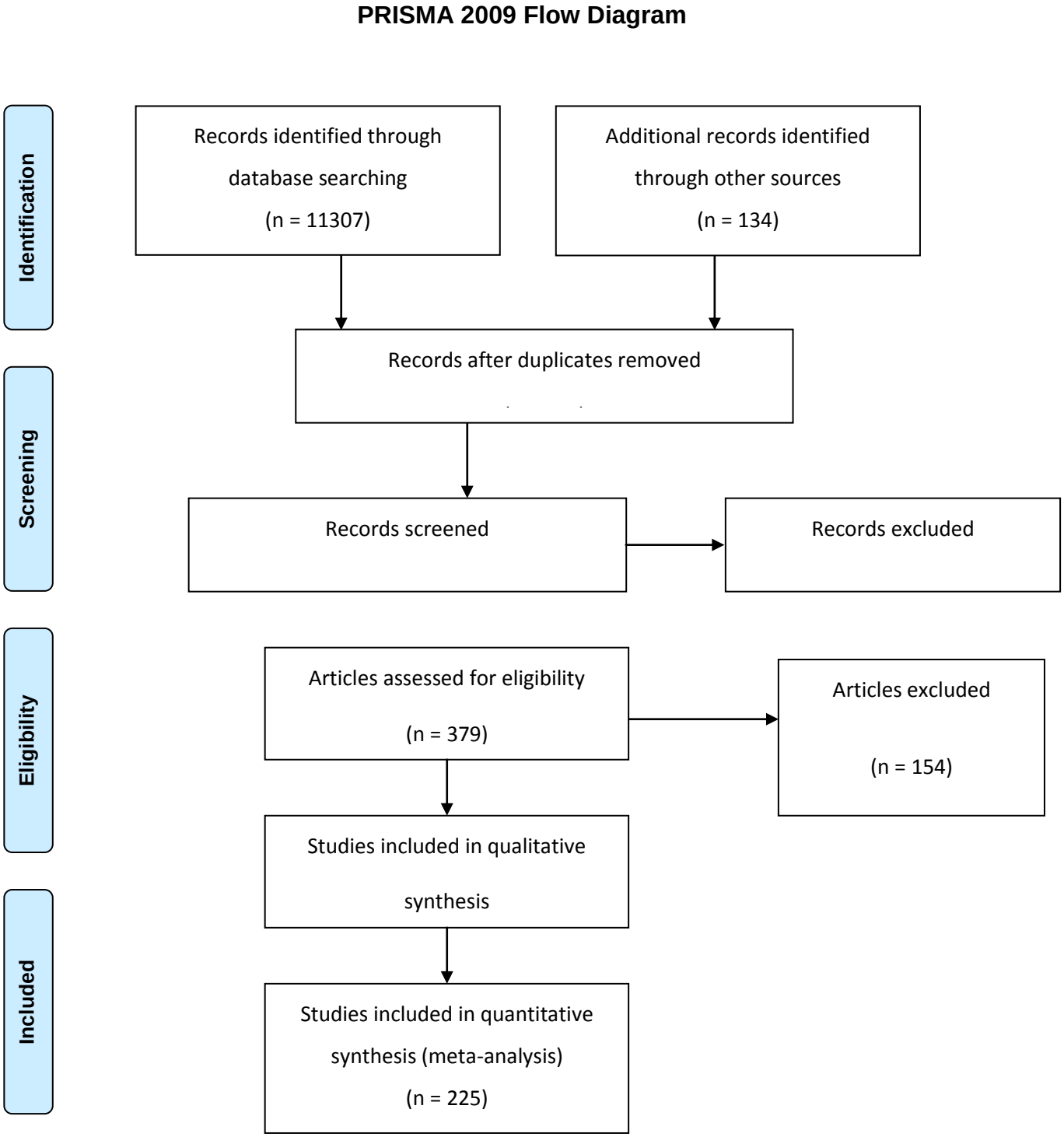
Appendix 7 – Population Frequency script

```
#!/usr/bin/env python3



def is_pos(pos_str):
    ''' if the string is like '21:123455' or 'X:123' etc '''
    parts=pos_str.split(':')
    if len(parts) != 2:
        return False
    if parts[0] not in ('1','2','3','4','5','6','7','8','9','10','11','12','13',
          '14','15','16','17','18','19','20','21','22','X'):
        return False
    if parts[1].isdigit():
        return True
    return False



#########################################################################
# read the list of SNP positions from the input table
snp_pos=set([])
ifile=open('population_frequency.txt')
for line in ifile:
    cols=line.split('\t')
    if is_pos(cols[6]):
```

```
        snp_pos.add(cols[6])
    else:
        continue
ifile.close()
# the input table is a bit missformatted on some lines (447,480)
# for example: line 480
# one extra tab after 'NP_001136.1:p.Phe100Ile'.
# The position string is now in column 8, instead of column 7.
# Removed it manually.
#########################################################################
# for all files in directory 1kg
# read the SNP lines, ignore the indels.

import os
os.chdir('1kg')
snp_lines={}
for pos in snp_pos:
    snp_lines[pos]=''

for filename in os.listdir('.'):
    ifile=open(filename)
    for line in ifile:
```

```
        cols=line.split('\t')
        pos_str=cols[0]
        if pos_str in snp_pos:
            if len(cols[1])==1 and len(cols[2])==1: # SNP
                snp_lines[pos_str]=snp_lines[pos_str]+'\t'+line[:-1]
            else: # indels
                continue
    ifile.close()


os.chdir('..')


################################################################################
####
# re-open the input table, appending the lines from 1kg
ifile=open('population_frequency.txt')
for line in ifile:
    cols=line.split('\t')
    if cols[6] not in snp_pos:
        print(line[:-1])
    else:
        print(line[:-1]+'\t'+snp_lines[cols[6]])
ifile.close()
```

Appendix 8 - Data extraction from publications

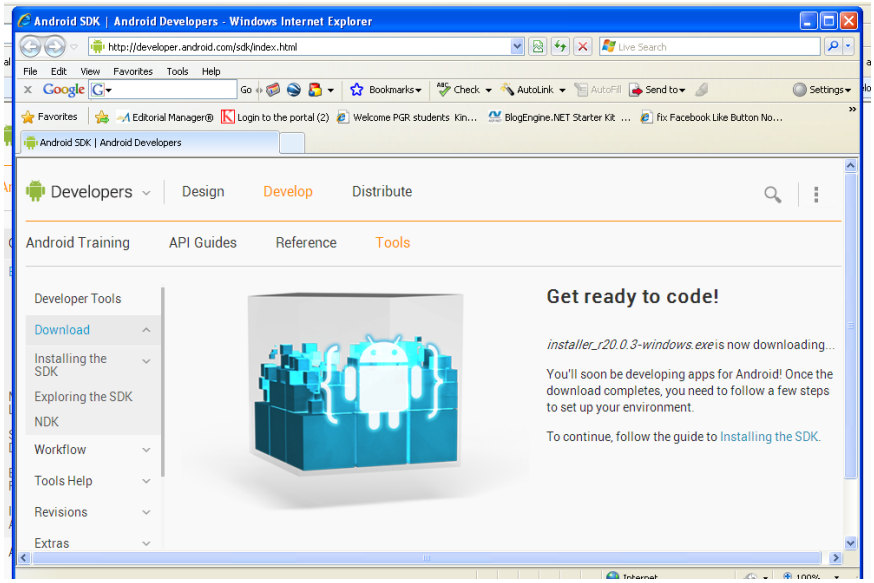


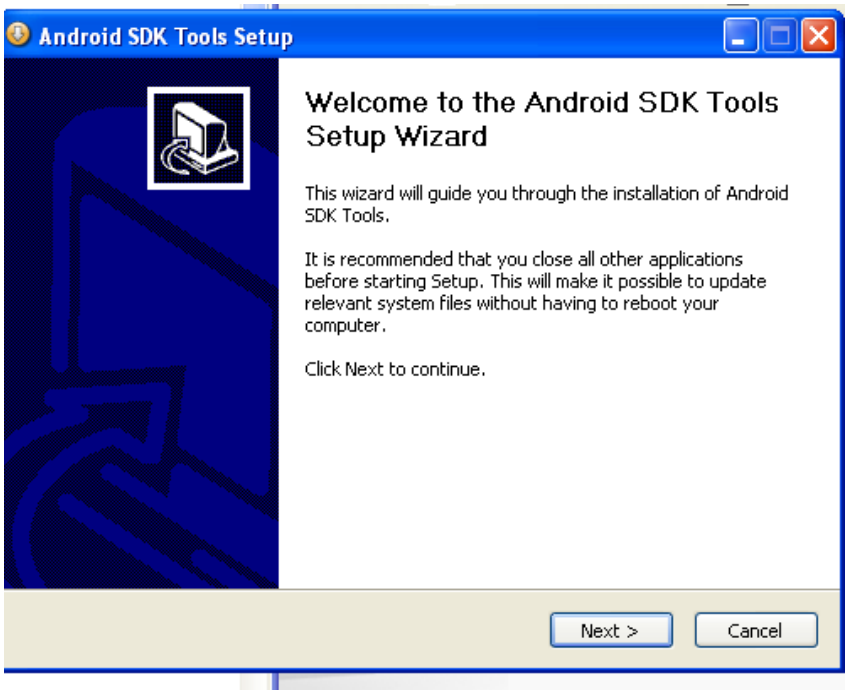
Appendix 9 – Population Frequency script using python

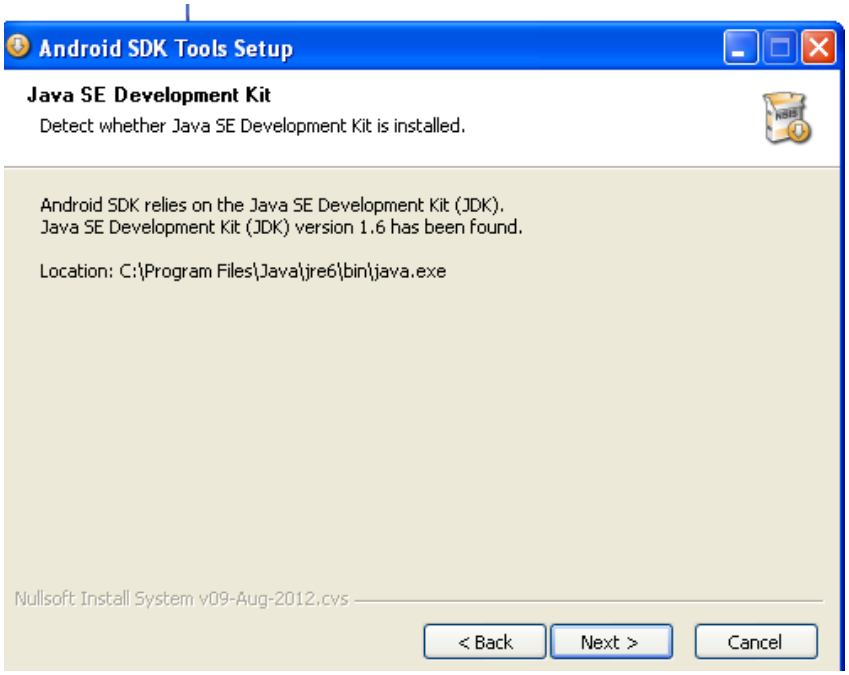
```
#!/usr/bin/env python3
def is_pos(pos_str):
    ''' if the string is like '21:123455' or 'X:123' etc '''
    parts=pos_str.split(':')
    if len(parts) != 2:
        return False
    if parts[0] not in ('1','2','3','4','5','6','7','8','9','10','11','12','13','14','15','16','17','18','19','20','21','22','X'):
        return False
    if parts[1].isdigit():
        return True
    return False

##############################################################################################
# read the list of SNP positions from the input table

snp_pos=set([])
ifile=open('population_frequency.txt')
for line in ifile:
    cols=line.split('\t')
    if is_pos(cols[6]):
        snp_pos.add(cols[6])
    else:
        continue
ifile.close()
# the input table is a bit missformatted on some lines (447,480)
# for example: line 480
# one extra tab after 'NP_001136.1:p.Phe100Ile'.
# The position string is now in column 8, instead of column 7.
# Removed it manually.

##############################################################################################
# for all files in directory 1kg
```

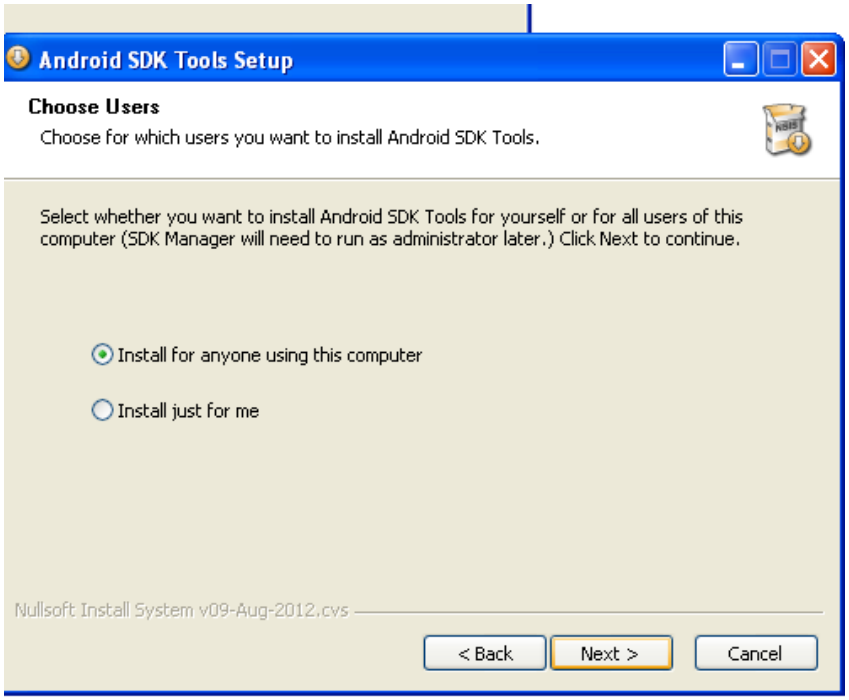
```
# read the SNP lines, ignore the indels.
import os
os.chdir('1kg')

snp_lines={}
for pos in snp_pos:
    snp_lines[pos]=''

for filename in os.listdir('.'):
    ifile=open(filename)
    for line in ifile:
        cols=line.split('\t')
        pos_str=cols[0]
        if pos_str in snp_pos:
            if len(cols[1])==1 and len(cols[2])==1: # SNP
                snp_lines[pos_str]=snp_lines[pos_str]+'\t'+line[:-1]
            else: # indels
                continue
    ifile.close()
os.chdir('..')
################################################################################
# re-open the input table, appending the lines from 1kg
ifile=open('population_frequency.txt')
for line in ifile:
    cols=line.split('\t')
    if cols[6] not in snp_pos:
        print(line[:-1])
    else:
        print(line[:-1]+'\t'+snp_lines[cols[6]])

ifile.close()
```

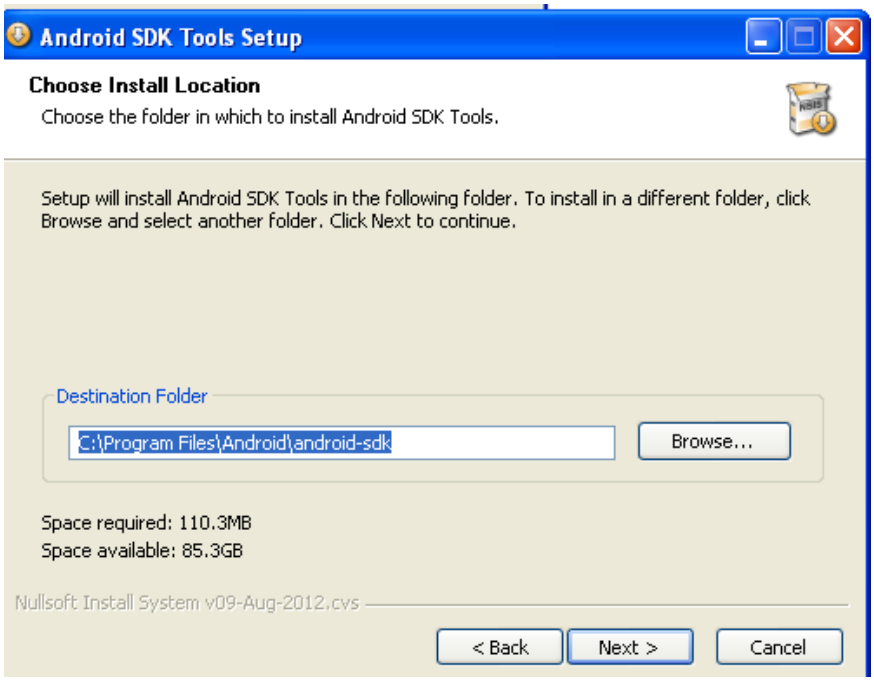
Appendix 10 - Process of installing and executing Eclipse application

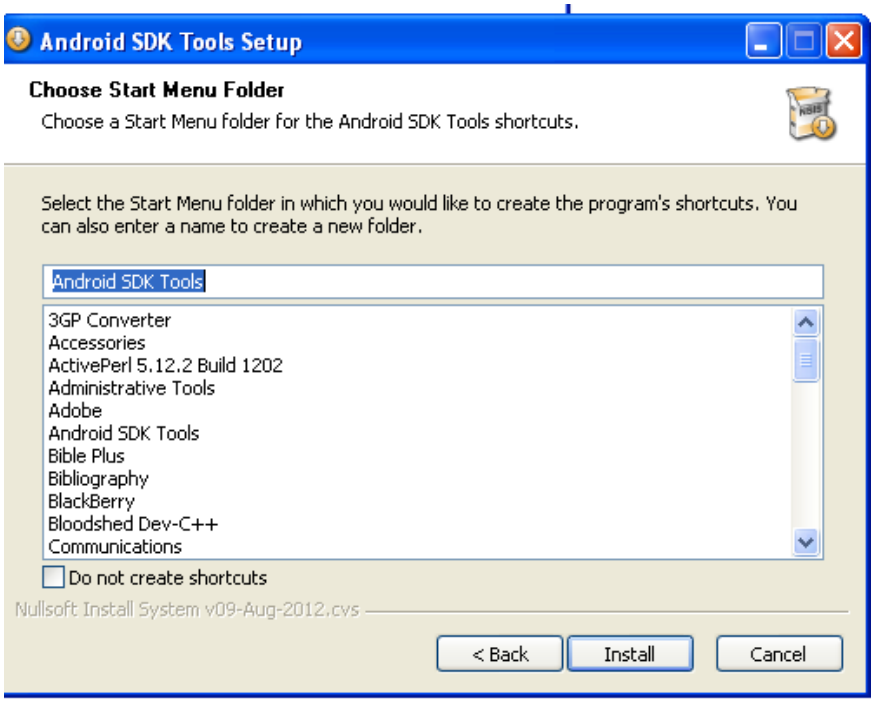


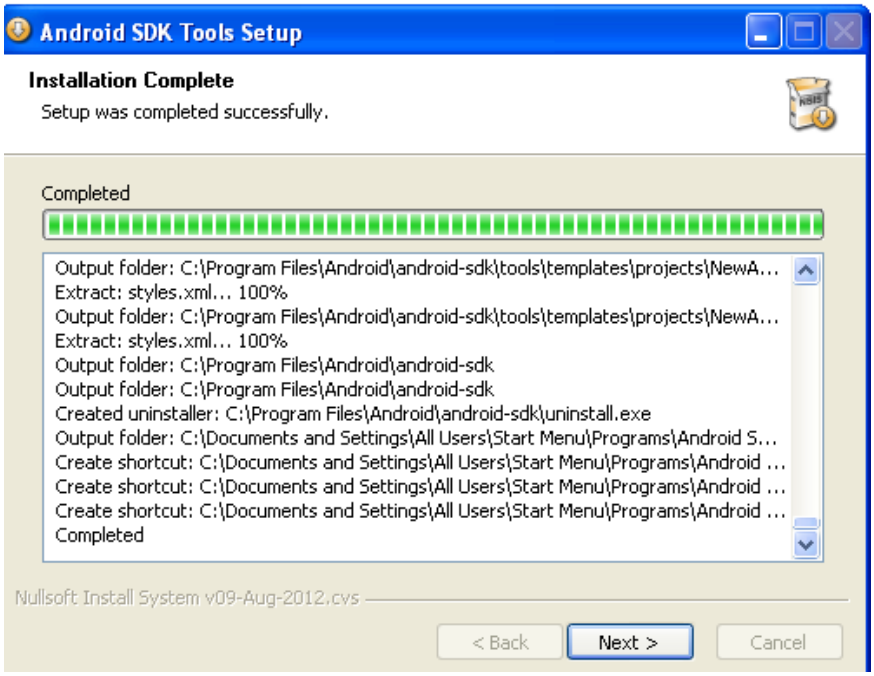


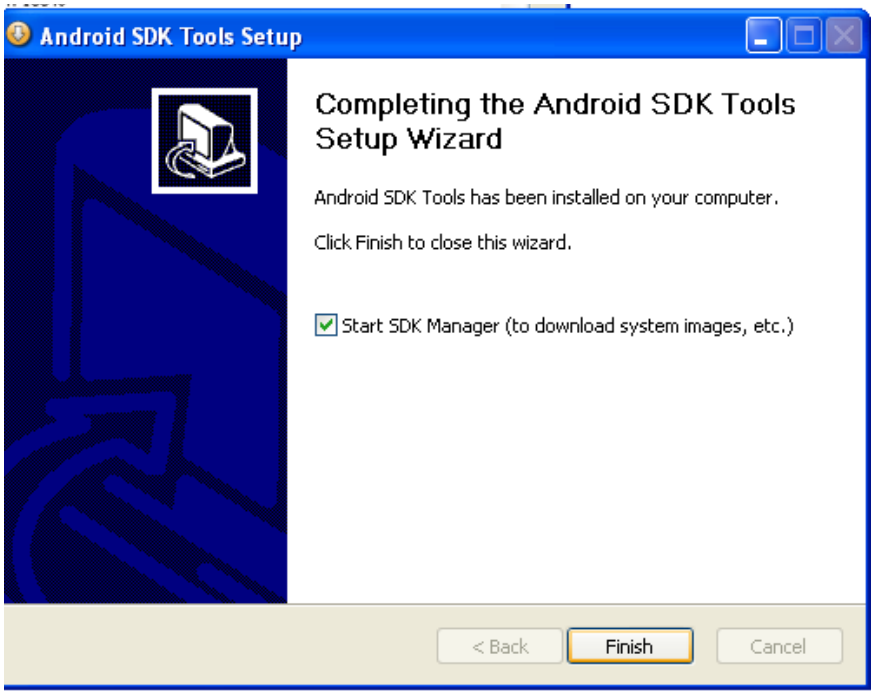

Android SDK Tools Setup
Java SE Development Kit
Detect whether Java SE Development Kit is installed.
Android SDK relies on the Java SE Development Kit (JDK).
Java SE Development Kit (JDK) version 1.6 has been found.
Location: C:\Program Files\Java\jre6\bin\java.exe
Nullsoft Install System v09-Aug-2012.cvs
< Back
Next >
Cancel

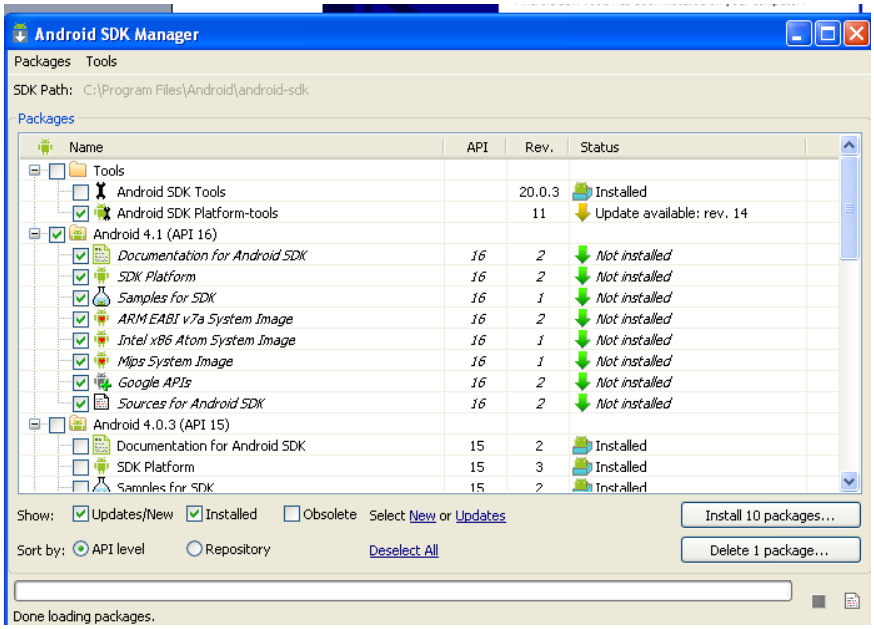

Android SDK Tools Setup
Choose Users
Choose for which users you want to install Android SDK Tools.
Select whether you want to install Android SDK Tools for yourself or for all users of this computer (SDK Manager will need to run as administrator later.) Click Next to continue.
Install for anyone using this computer
Install just for me
Nullsoft Install System v09-Aug-2012.cvs
< Back
Next >
Cancel

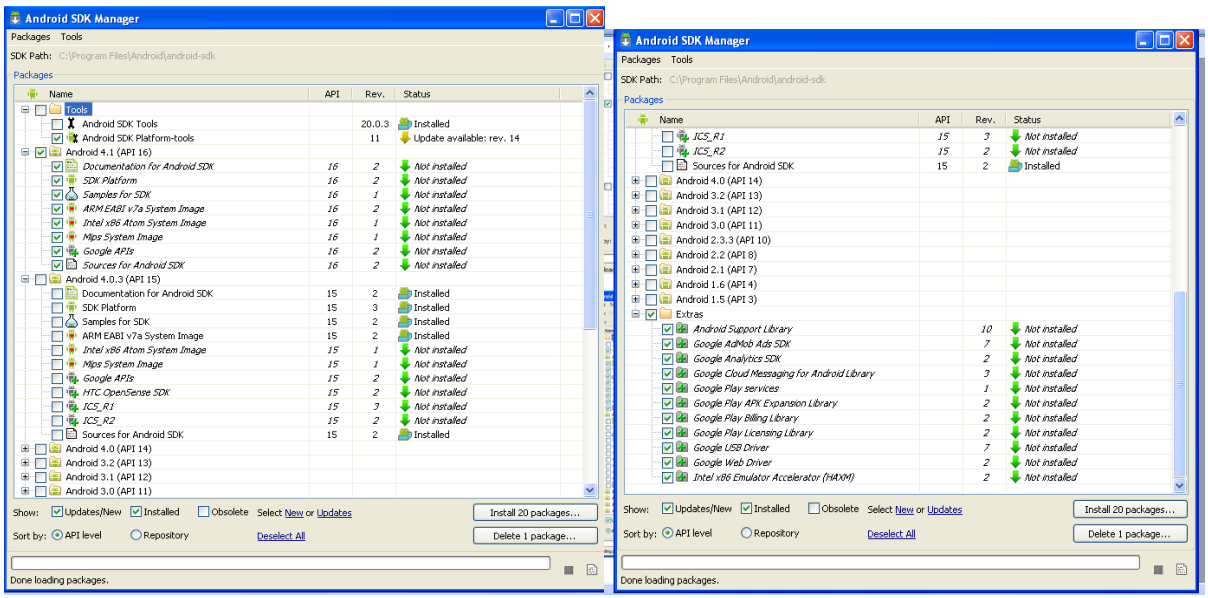

Android SDK Tools Setup
Choose Install Location
Choose the folder in which to install Android SDK Tools.
Setup will install Android SDK Tools in the following folder. To install in a different folder, click Browse and select another folder. Click Next to continue.
Destination Folder
C:\Program Files\Android\android-sdk
Browse...
Space required: 110.3MB
Space available: 85.3GB
Nullsoft Install System v09-Aug-2012.cvs
< Back
Next >
Cancel

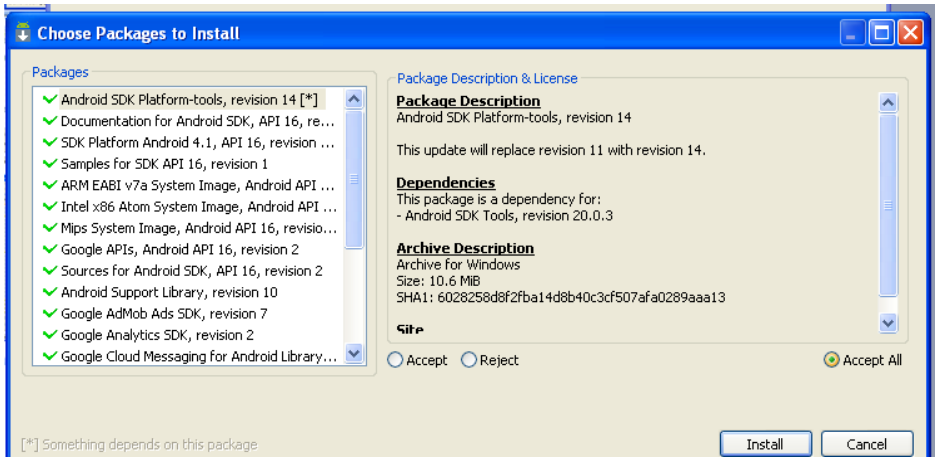

Android SDK Tools Setup
Choose Start Menu Folder
Choose a Start Menu folder for the Android SDK Tools shortcuts.
Select the Start Menu folder in which you would like to create the program's shortcuts. You can also enter a name to create a new folder.
Android SDK Tools
3GP Converter
Accessories
ActivePerl 5.12.2 Build 1202
Administrative Tools
Adobe
Android SDK Tools
Bible Plus
Bibliography
BlackBerry
Bloodshed Dev-C++
Communications
Do not create shortcuts
Nullsoft Install System v09-Aug-2012.cvs
< Back
Install
Cancel

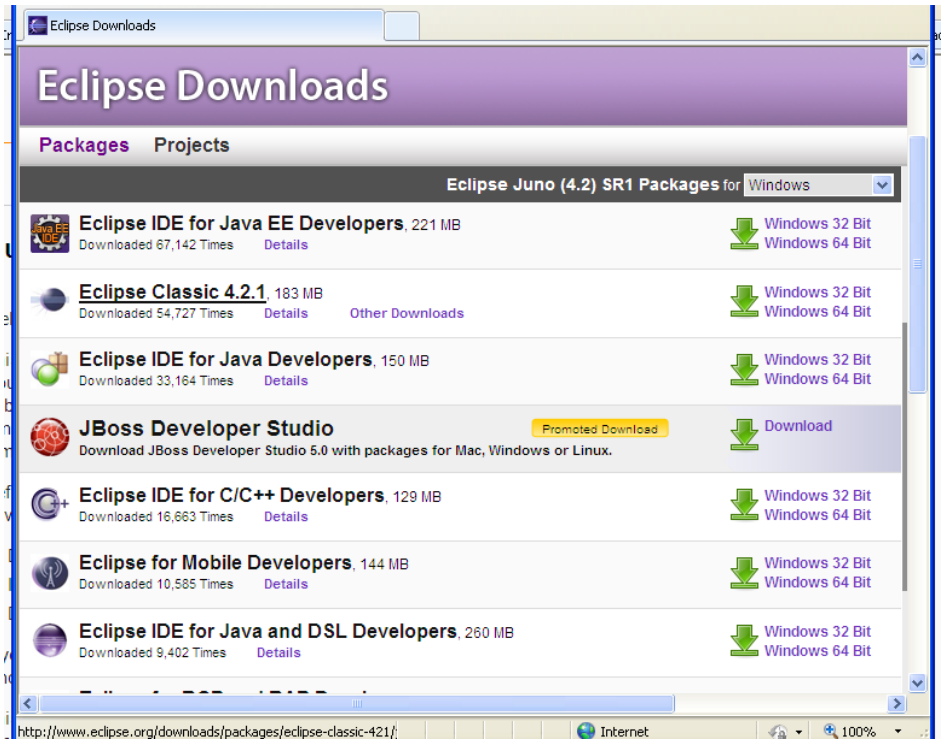

Android SDK Tools Setup
Installation Complete
Setup was completed successfully.
Completed
Output folder: C:\Program Files\Android\android-sdk\tools\templates\projects\NewA...
Extract: styles.xml... 100%
Output folder: C:\Program Files\Android\android-sdk\tools\templates\projects\NewA...
Extract: styles.xml... 100%
Output folder: C:\Program Files\Android\android-sdk
Output folder: C:\Program Files\Android\android-sdk
Created uninstaller: C:\Program Files\Android\android-sdk\uninstall.exe
Output folder: C:\Documents and Settings\All Users\Start Menu\Programs\Android S...
Create shortcut: C:\Documents and Settings\All Users\Start Menu\Programs\Android ...
Create shortcut: C:\Documents and Settings\All Users\Start Menu\Programs\Android ...
Create shortcut: C:\Documents and Settings\All Users\Start Menu\Programs\Android ...
Completed
Nullsoft Install System v09-Aug-2012.cvs
< Back
Next >
Cancel

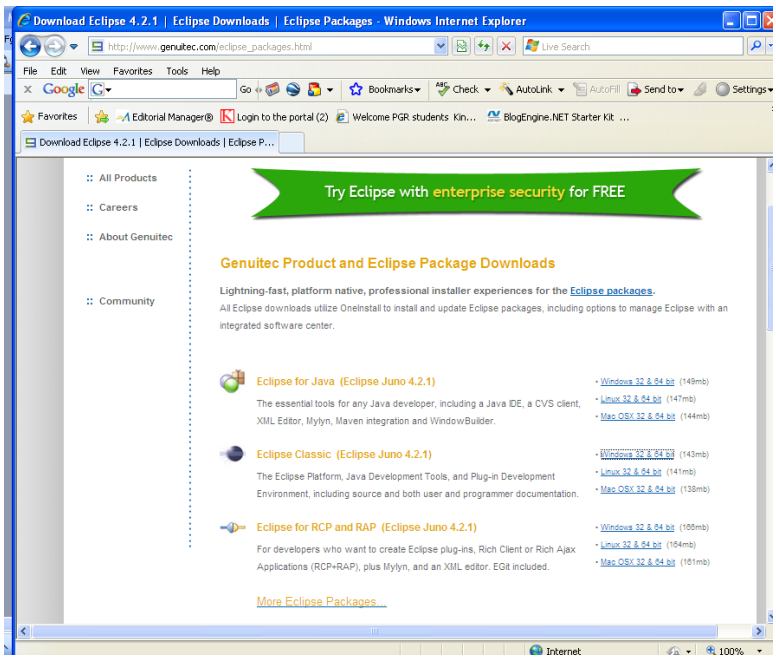

Android SDK Tools Setup
Completing the Android SDK Tools Setup Wizard
Android SDK Tools has been installed on your computer.
Click Finish to close this wizard.
Start SDK Manager (to download system images, etc.)
< Back
Finish
Cancel

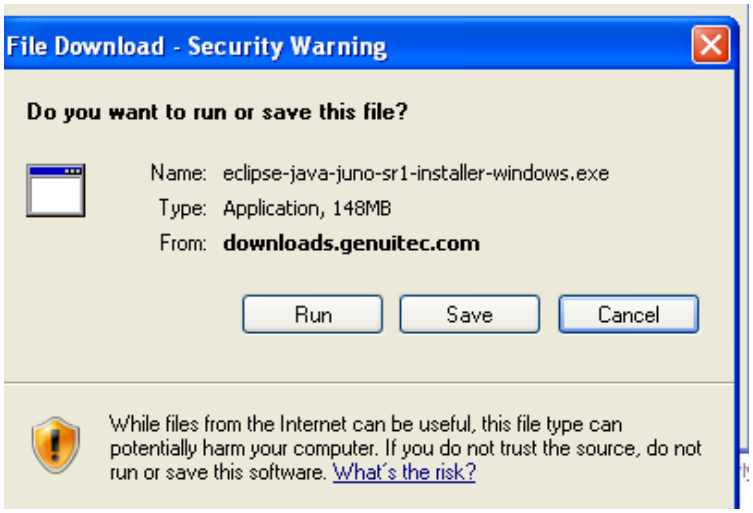

Android SDK Manager
Packages Tools
SDK Path: C:\Program Files\Android\android-sdk
Packages
Name API Rev. Status
Tools
Android SDK Tools 20.0.3 Installed
Android SDK Platform-tools 11 Update available: rev. 14
Android 4.1 (API 16)
Documentation for Android SDK 16 2 Not installed
SDK Platform 16 2 Not installed
Samples for SDK 16 1 Not installed
ARM EABI v7a System Image 16 2 Not installed
Intel x86 Atom System Image 16 1 Not installed
Mips System Image 16 1 Not installed
Google APIs 16 2 Not installed
Sources for Android SDK 16 2 Not installed
Android 4.0.3 (API 15)
Documentation for Android SDK 15 2 Installed
SDK Platform 15 3 Installed
Samples for SDK 15 2 Installed
Show: Updates/New Installed Obsolete Select New or Updates
Install 10 packages...
Sort by: API level Repository
Deselect All
Delete 1 package...
Done loading packages.

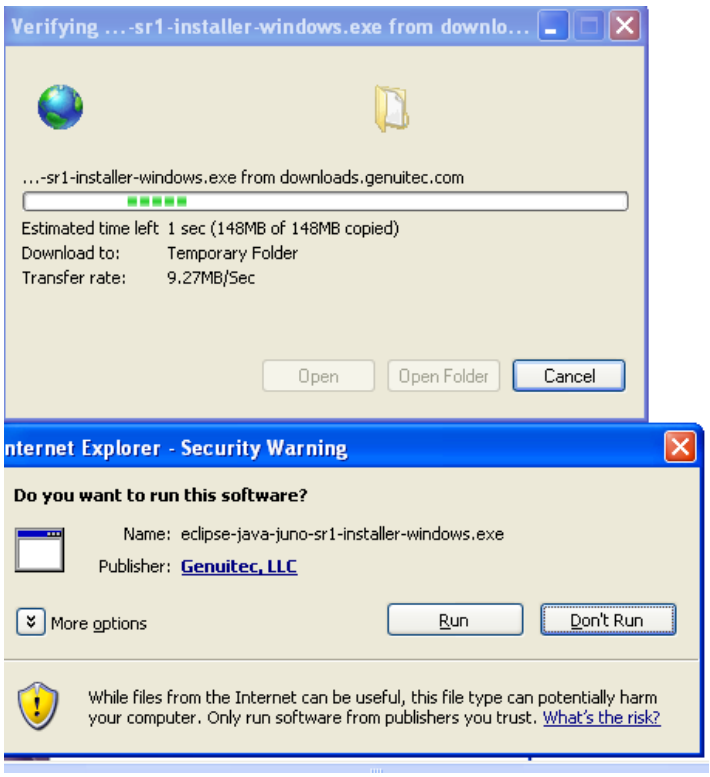

Android SDK Manager
Packages Tools
SDK Path: C:\Program Files\Android\android-sdk
Packages
Name API Rev. Status
Tools
Android SDK Tools 20.0.3 Installed
Android SDK Platform-tools 11 Update available: rev. 14
Android 4.1 (API 16)
Documentation for Android SDK 16 2 Not installed
SDK Platform 16 2 Not installed
Samples for SDK 16 1 Not installed
ARM EABI v7a System Image 16 2 Not installed
Intel x86 Atom System Image 16 1 Not installed
Mips System Image 16 1 Not installed
Google APIs 16 2 Not installed
Sources for Android SDK 16 2 Not installed
Android 4.0.3 (API 15)
Documentation for Android SDK 15 2 Installed
SDK Platform 15 3 Installed
Samples for SDK 15 2 Installed
ARM EABI v7a System Image 15 2 Installed
Intel x86 Atom System Image 15 1 Not installed
Mips System Image 15 1 Not installed
Google APIs 15 2 Not installed
HTC OpenSense SDK 15 2 Not installed
ICS_R1 15 3 Not installed
ICS_R2 15 2 Not installed
Sources for Android SDK 15 2 Installed
Android 4.0 (API 14)
Android 3.2 (API 13)
Android 3.1 (API 12)
Android 3.0 (API 11)
Show: Updates/New Installed Obsolete Select New or Updates
Install 20 packages...
Sort by: API level Repository
Deselect All
Delete 1 package...
Done loading packages.
Android SDK Manager
Packages Tools
SDK Path: C:\Program Files\Android\android-sdk
Packages
Name API Rev. Status
ICS_R1 15 3 Not installed
ICS_R2 15 2 Not installed
Sources for Android SDK 15 2 Installed
Android 4.0 (API 14)
Android 3.2 (API 13)
Android 3.1 (API 12)
Android 3.0 (API 11)
Android 2.3.3 (API 10)
Android 2.2 (API 8)
Android 2.1 (API 7)
Android 1.6 (API 4)
Android 1.5 (API 3)
Extras
Android Support Library 10 Not installed
Google AdMob Ads SDK 7 Not installed
Google Analytics SDK 2 Not installed
Google Cloud Messaging for Android Library 3 Not installed
Google Play services 1 Not installed
Google Play APK Expansion Library 2 Not installed
Google Play Billing Library 2 Not installed
Google Play Licensing Library 2 Not installed
Google USB Driver 7 Not installed
Google Web Driver 2 Not installed
Intel x86 Emulator Accelerator (HAXM) 2 Not installed
Show: Updates/New Installed Obsolete Select New or Updates
Install 20 packages...
Sort by: API level Repository
Deselect All
Delete 1 package...
Done loading packages.

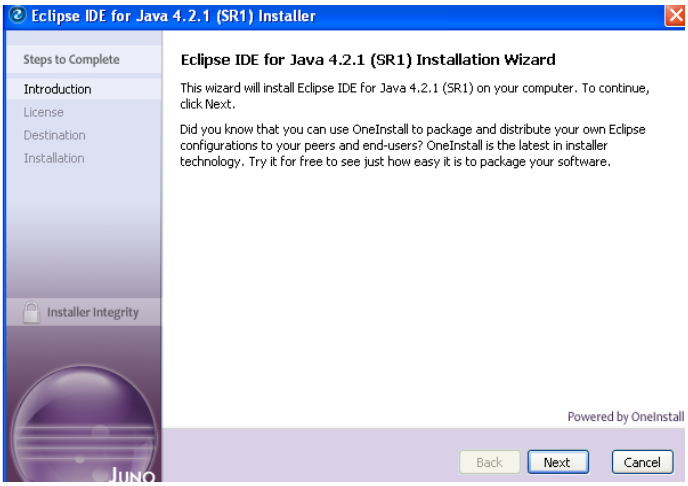

Choose Packages to Install
Packages
Android SDK Platform-tools, revision 14 [*]
Documentation for Android SDK, API 16, re...
SDK Platform Android 4.1, API 16, revision ...
Samples for SDK API 16, revision 1
ARM EABI v7a System Image, Android API ...
Intel x86 Atom System Image, Android API ...
Mips System Image, Android API 16, revisio...
Google APIs, Android API 16, revision 2
Sources for Android SDK, API 16, revision 2
Android Support Library, revision 10
Google AdMob Ads SDK, revision 7
Google Analytics SDK, revision 2
Google Cloud Messaging for Android Library...
Package Description & License
Package Description
Android SDK Platform-tools, revision 14
This update will replace revision 11 with revision 14.
Dependencies
This package is a dependency for:
- Android SDK Tools, revision 20.0.3
Archive Description
Archive for Windows
Size: 10.6 MiB
SHA1: 6028258d8f2fba14d8b40c3cf507afa0289aaa13
Site
Accept
Reject
Accept All
[*] Something depends on this package
Install
Cancel

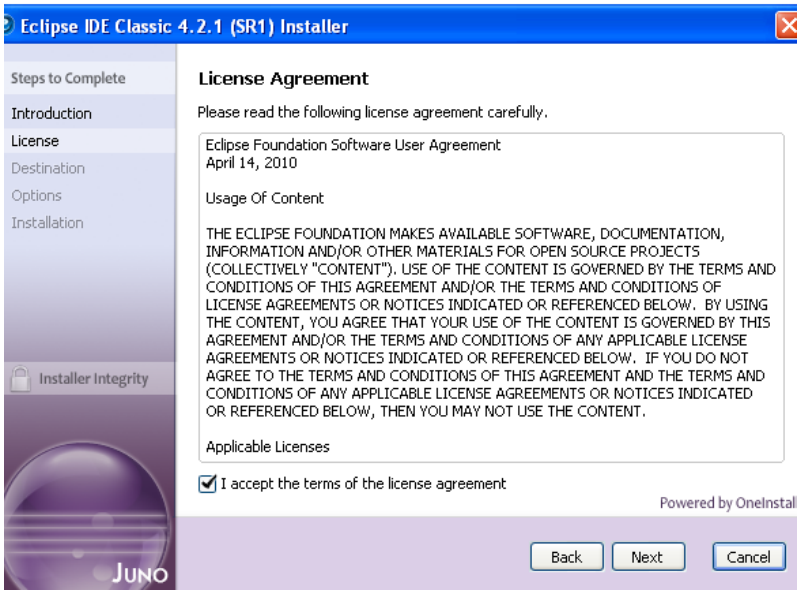

Eclipse Downloads
Eclipse Downloads
Packages
Projects
Eclipse Juno (4.2) SR1 Packages for Windows
Eclipse IDE for Java EE Developers, 221 MB
Downloaded 67,142 Times
Details
Windows 32 Bit
Windows 64 Bit
Eclipse Classic 4.2.1, 183 MB
Downloaded 54,727 Times
Details
Other Downloads
Windows 32 Bit
Windows 64 Bit
Eclipse IDE for Java Developers, 150 MB
Downloaded 33,164 Times
Details
Windows 32 Bit
Windows 64 Bit
JBoss Developer Studio
Promoted Download
Download JBoss Developer Studio 5.0 with packages for Mac, Windows or Linux.
Download
Eclipse IDE for C/C++ Developers, 129 MB
Downloaded 16,663 Times
Details
Windows 32 Bit
Windows 64 Bit
Eclipse for Mobile Developers, 144 MB
Downloaded 10,585 Times
Details
Windows 32 Bit
Windows 64 Bit
Eclipse IDE for Java and DSL Developers, 260 MB
Downloaded 9,402 Times
Details
Windows 32 Bit
Windows 64 Bit
http://www.eclipse.org/downloads/packages/eclipse-classic-421/
Internet
100%

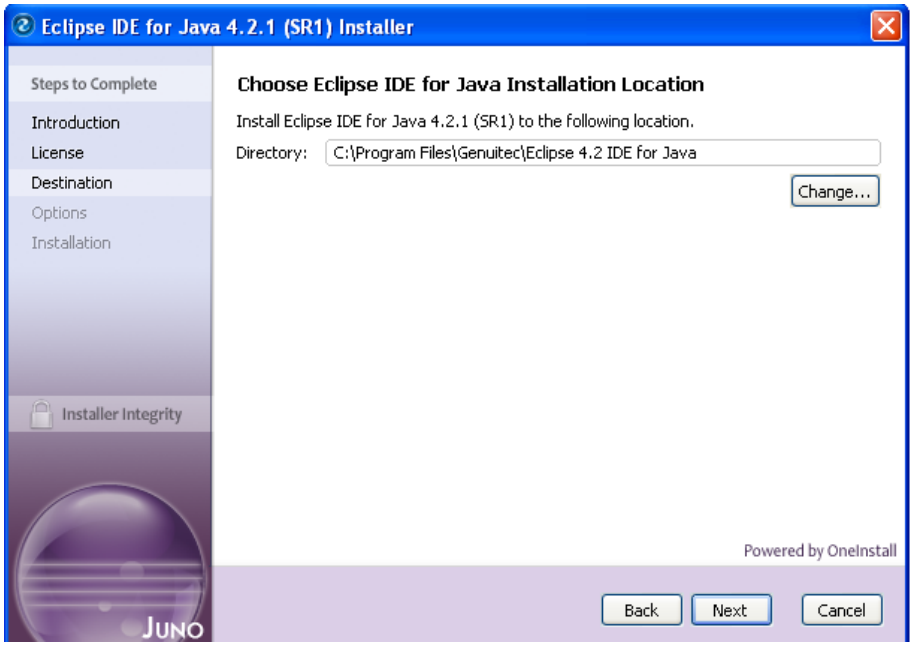

Download Eclipse 4.2.1 | Eclipse Downloads | Eclipse Packages - Windows Internet Explorer
http://www.genuitec.com/eclipse_packages.html
Try Eclipse with enterprise security for FREE
All Products
Careers
About Genuitec
Community
Genuitec Product and Eclipse Package Downloads
Lightning-fast, platform native, professional installer experiences for the Eclipse packages.
All Eclipse downloads utilize OneInstall to install and update Eclipse packages, including options to manage Eclipse with an integrated software center.
Eclipse for Java (Eclipse Juno 4.2.1)
The essential tools for any Java developer, including a Java IDE, a CVS client, XML Editor, Mylyn, Maven integration and WindowBuilder.
Eclipse Classic (Eclipse Juno 4.2.1)
The Eclipse Platform, Java Development Tools, and Plug-in Development Environment, including source and both user and programmer documentation.
Eclipse for RCP and RAP (Eclipse Juno 4.2.1)
For developers who want to create Eclipse plug-ins, Rich Client or Rich Ajax Applications (RCP+RAP), plus Mylyn, and an XML editor. EGit included.
More Eclipse Packages...

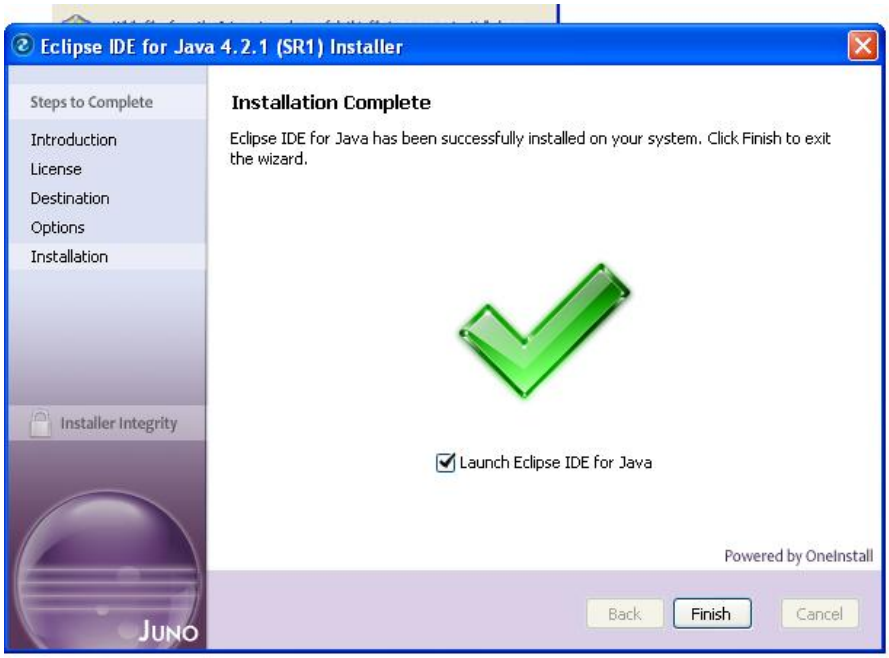

File Download - Security Warning
Do you want to run or save this file?
Name: eclipse-java-juno-sr1-installer-windows.exe
Type: Application, 148MB
From: downloads.genuitec.com
Run
Save
Cancel
While files from the Internet can be useful, this file type can potentially harm your computer. If you do not trust the source, do not run or save this software. What's the risk?

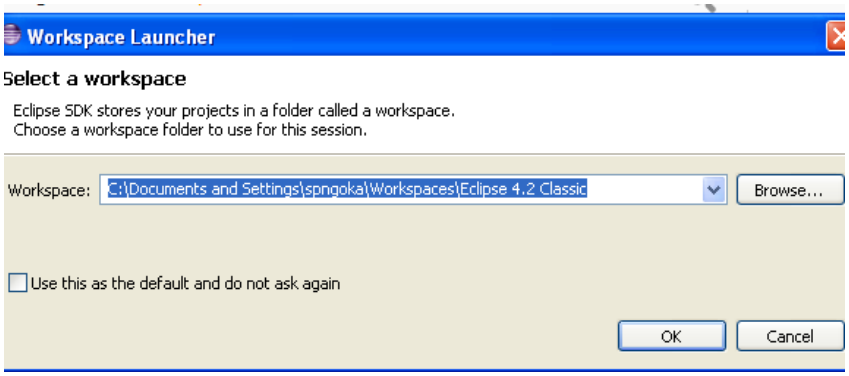

Verifying ...-sr1-installer-windows.exe from downlo...
...-sr1-installer-windows.exe from downloads.genuitec.com
Estimated time left 1 sec (148MB of 148MB copied)
Download to: Temporary Folder
Transfer rate: 9.27MB/Sec
Open
Open Folder
Cancel
nternet Explorer - Security Warning
Do you want to run this software?
Name: eclipse-java-juno-sr1-installer-windows.exe
Publisher: Genuitec, LLC
More options
Run
Don't Run
While files from the Internet can be useful, this file type can potentially harm your computer. Only run software from publishers you trust. What's the risk?

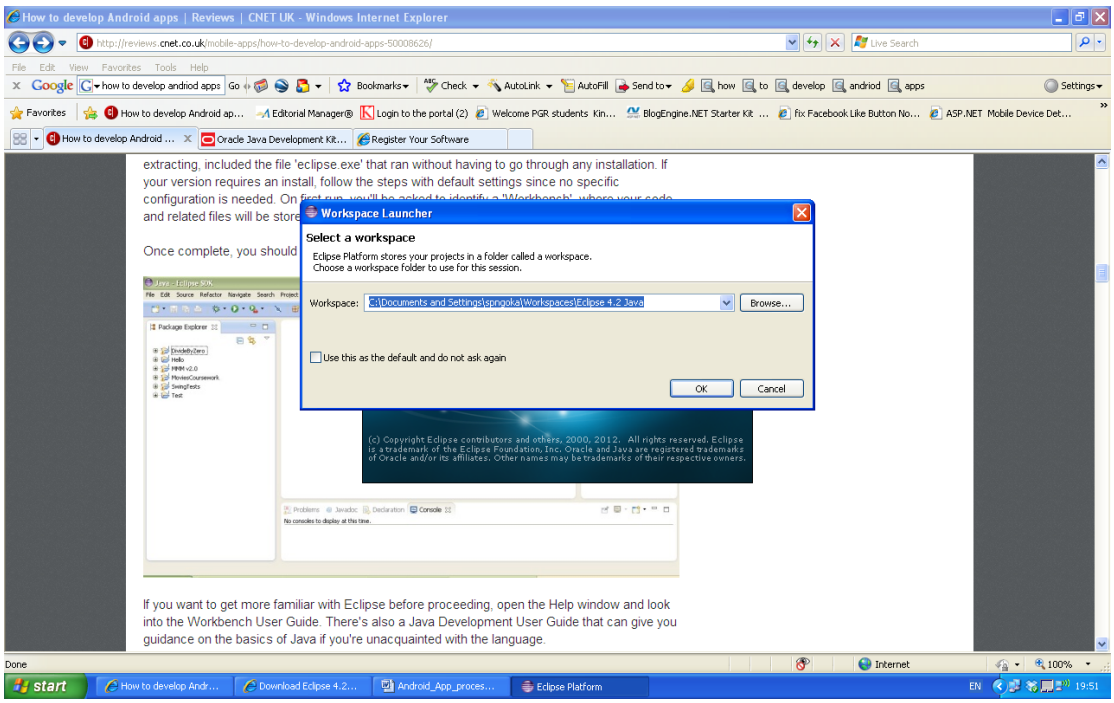

Eclipse IDE for Java 4.2.1 (SR1) Installer
Steps to Complete
Introduction
License
Destination
Installation
Installer Integrity
Eclipse IDE for Java 4.2.1 (SR1) Installation Wizard
This wizard will install Eclipse IDE for Java 4.2.1 (SR1) on your computer. To continue, click Next.
Did you know that you can use OneInstall to package and distribute your own Eclipse configurations to your peers and end-users? OneInstall is the latest in installer technology. Try it for free to see just how easy it is to package your software.
Powered by OneInstall
Back
Next
Cancel


Eclipse IDE Classic 4.2.1 (SR1) Installer
Steps to Complete
Introduction
License
Destination
Options
Installation
Installer Integrity
JUNO
License Agreement
Please read the following license agreement carefully.
Eclipse Foundation Software User Agreement
April 14, 2010
Usage Of Content
THE ECLIPSE FOUNDATION MAKES AVAILABLE SOFTWARE, DOCUMENTATION, INFORMATION AND/OR OTHER MATERIALS FOR OPEN SOURCE PROJECTS (COLLECTIVELY "CONTENT"). USE OF THE CONTENT IS GOVERNED BY THE TERMS AND CONDITIONS OF THIS AGREEMENT AND/OR THE TERMS AND CONDITIONS OF LICENSE AGREEMENTS OR NOTICES INDICATED OR REFERENCED BELOW. BY USING THE CONTENT, YOU AGREE THAT YOUR USE OF THE CONTENT IS GOVERNED BY THIS AGREEMENT AND/OR THE TERMS AND CONDITIONS OF ANY APPLICABLE LICENSE AGREEMENTS OR NOTICES INDICATED OR REFERENCED BELOW. IF YOU DO NOT AGREE TO THE TERMS AND CONDITIONS OF THIS AGREEMENT AND THE TERMS AND CONDITIONS OF ANY APPLICABLE LICENSE AGREEMENTS OR NOTICES INDICATED OR REFERENCED BELOW, THEN YOU MAY NOT USE THE CONTENT.
Applicable Licenses
I accept the terms of the license agreement
Powered by OneInstal
Back
Next
Cancel


Eclipse IDE for Java 4.2.1 (SR1) Installer
Steps to Complete
Introduction
License
Destination
Options
Installation
Installer Integrity
JUNO
Choose Eclipse IDE for Java Installation Location
Install Eclipse IDE for Java 4.2.1 (SR1) to the following location.
Directory:
C:\Program Files\Genuitec\Eclipse 4.2 IDE for Java
Change...
Powered by OneInstall
Back
Next
Cancel


Eclipse IDE for Java 4.2.1 (SR1) Installer
Steps to Complete
Introduction
License
Destination
Options
Installation
Installer Integrity
JUNO
Installation Complete
Eclipse IDE for Java has been successfully installed on your system. Click Finish to exit the wizard.
Launch Eclipse IDE for Java
Powered by OneInstall
Back
Finish
Cancel


Workspace Launcher
Select a workspace
Eclipse SDK stores your projects in a folder called a workspace.
Choose a workspace folder to use for this session.
Workspace:
C:\Documents and Settings\spngoka\Workspaces\Eclipse 4.2 Classic
Browse...
Use this as the default and do not ask again
OK
Cancel



Download the ADT Plugin

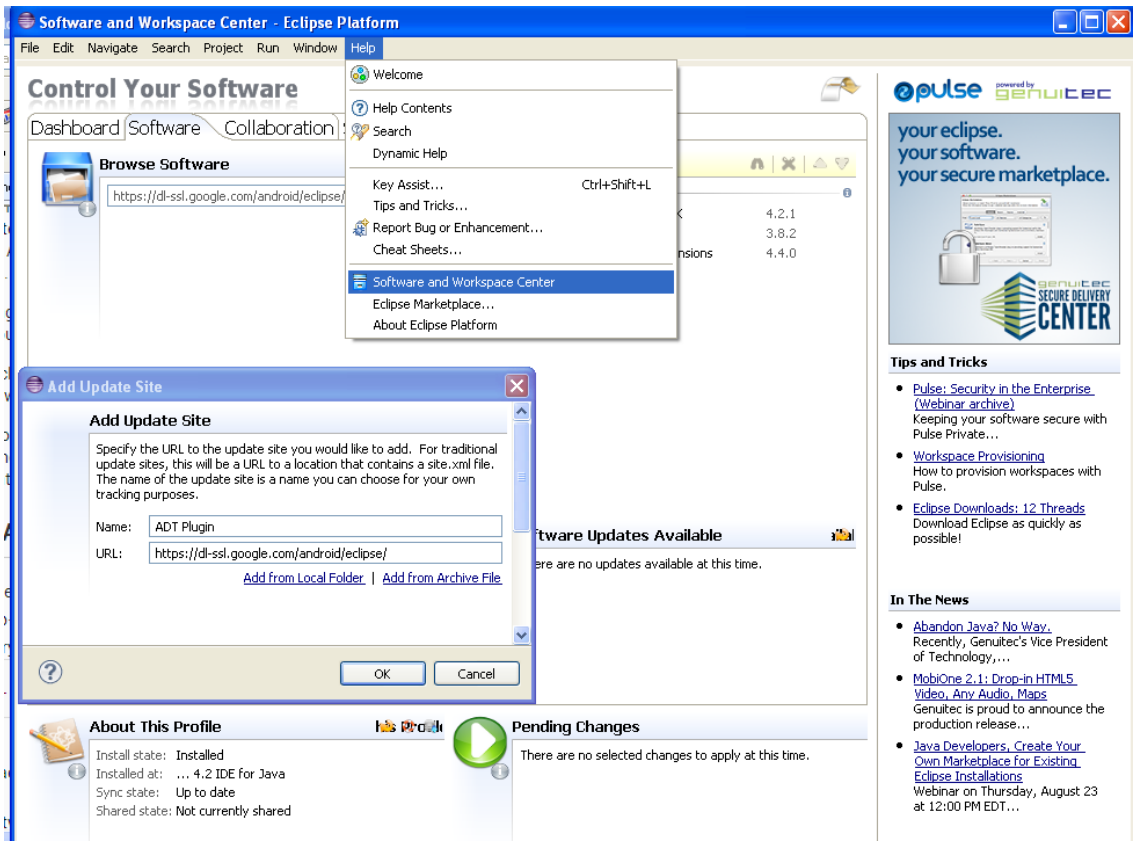
1. Start Eclipse, then select **Help > Install New Software**.
2. Click **Add**, in the top-right corner.
3. In the Add Repository dialog that appears, enter "ADT Plugin" for the *Name* and the following URL for the *Location*:

   ```
   https://dl-ssl.google.com/android/eclipse/
   ```

4. Click **OK**.

   If you have trouble acquiring the plugin, try using "http" in the Location URL, instead of "https" (https is preferred for security reasons).
5. In the Available Software dialog, select the checkbox next to Developer Tools and click **Next**.
6. In the next window, you'll see a list of the tools to be downloaded. Click **Next**.
7. Read and accept the license agreements, then click **Finish**.

   If you get a security warning saying that the authenticity or validity of the software can't be established, click **OK**.
8. When the installation completes, restart Eclipse.


Software and Workspace Center - Eclipse Platform
File Edit Navigate Search Project Run Window Help
Control Your Software
Dashboard Software Collaboration
Browse Software
Welcome
Help Contents
Search
Dynamic Help
Key Assist... Ctrl+Shift+L
Tips and Tricks...
Report Bug or Enhancement...
Cheat Sheets...
Software and Workspace Center
Eclipse Marketplace...
About Eclipse Platform
Add Update Site
Add Update Site
Specify the URL to the update site you would like to add. For traditional update sites, this will be a URL to a location that contains a site.xml file. The name of the update site is a name you can choose for your own tracking purposes.
Name: ADT Plugin
URL: https://dl-ssl.google.com/android/eclipse/
Add from Local Folder | Add from Archive File
OK Cancel
tware Updates Available
ere are no updates available at this time.
About This Profile
Install state: Installed
Installed at: ... 4.2 IDE for Java
Sync state: Up to date
Shared state: Not currently shared
Pending Changes
There are no selected changes to apply at this time.
your eclipse.
your software.
your secure marketplace.
SECURE DELIVERY CENTER
Tips and Tricks
Pulse: Security in the Enterprise (Webinar archive)
Keeping your software secure with Pulse Private...
Workspace Provisioning
How to provision workspaces with Pulse.
Eclipse Downloads: 12 Threads
Download Eclipse as quickly as possible!
In The News
Abandon Java? No Way.
Recently, Genuitec's Vice President of Technology,...
MobiOne 2.1: Drop-in HTML5 Video, Any Audio, Maps
Genuitec is proud to announce the production release...
Java Developers, Create Your Own Marketplace for Existing Eclipse Installations
Webinar on Thursday, August 23 at 12:00 PM EDT...


Eclipse
A Java Runtime Environment (JRE) or Java Development Kit (JDK) must be available in order to run Eclipse. No Java virtual machine was found after searching the following locations:
C:\Program Files\Genuitec\Eclipse 4.2 IDE for Java\jre\bin\javaw.exe
javaw.exe in your current PATH
OK


Eclipse: eclipse.exe - Unable To Locate Component
This application has failed to start because verify.dll was not found. Re-installing the application may fix this problem.
OK

Configure the ADT Plugin

Once Eclipse restarts, you must specify the location of your Android SDK directory:

1. In the "Welcome to Android Development" window that appears, select **Use existing SDKs**.
2. Browse and select the location of the Android SDK directory you recently downloaded.
3. Click **Next**.

If you haven't encountered any errors, you're done setting up ADT and can continue to Next Steps.

Updating the ADT Plugin

From time to time, a new revision of the ADT Plugin becomes available, with new features and bug fixes. Generally, when a new revision of ADT is available, you should update to it as soon as convenient.

In some cases, a new revision of ADT will have a dependency on a specific revision of the Android SDK Tools. If such dependencies exist, you will need to update the SDK Tools package of the SDK after installing the new revision of ADT. To update the SDK Tools package, use the Android SDK Manager, as described in Exploring the SDK.

To learn about new features of each ADT revision and also any dependencies on the SDK Tools, see the listings in the Revisions section. To determine the version currently installed, open the Eclipse Installed Software window using **Help** > **Software Updates** and refer to the version listed for "Android Development Tools".

Follow the steps below to check whether an update is available and, if so, to install it.

1. Select **Help** > **Check for Updates**.
   If there are no updates available, a dialog will say so and you're done.
2. If there are updates available, select Android DDMS, Android Development Tools, and Android Hierarchy Viewer, then click **Next**.
3. In the Update Details dialog, click **Next**.
4. Read and accept the license agreement and then click **Finish**. This will download and install the latest version of Android DDMS and Android Development Tools.
5. Restart Eclipse.

If you encounter problems during the update, remove the existing ADT plugin from Eclipse, then perform a fresh installation, using the instructions for Installing the ADT Plugin.

Troubleshooting

If you are having trouble downloading the ADT plugin after following the steps above, here are some suggestions:

- If Eclipse can not find the remote update site containing the ADT plugin, try changing the remote site URL to use http, rather than https. That is, set the Location for the remote site to:

```
http://dl-ssl.google.com/android/eclipse/
```

- If you are behind a firewall (such as a corporate firewall), make sure that you have properly configured your proxy settings in Eclipse. In Eclipse, you can configure proxy information from the main Eclipse menu in **Window** (on Mac OS X, **Eclipse**) > **Preferences** > **General** > **Network Connections**.

If you are still unable to use Eclipse to download the ADT plugin as a remote update site, you can download the ADT zip file to your local machine and manually install it:

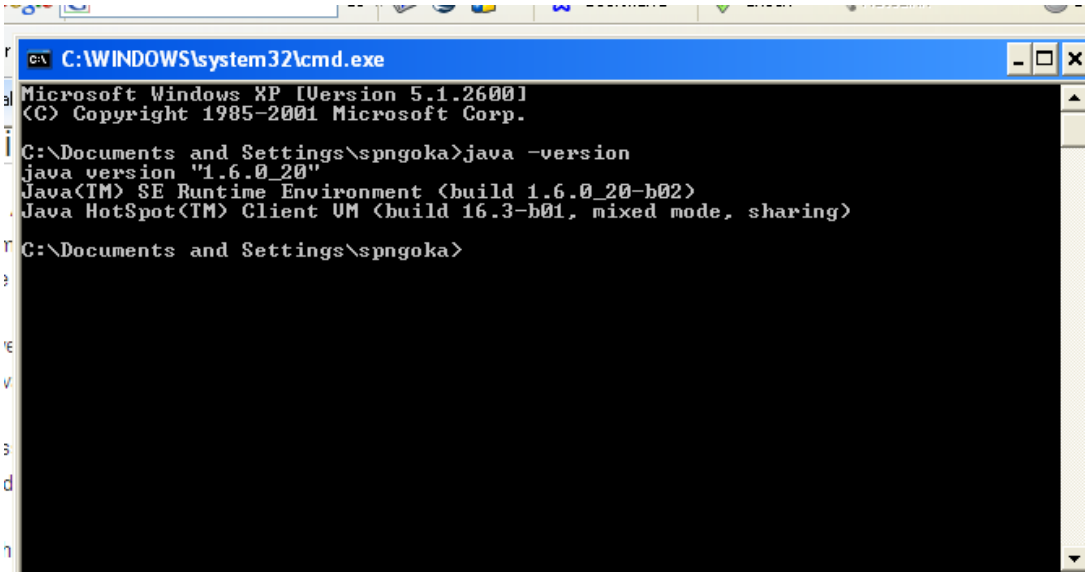
1. Download the ADT Plugin zip file (do not unpack it):

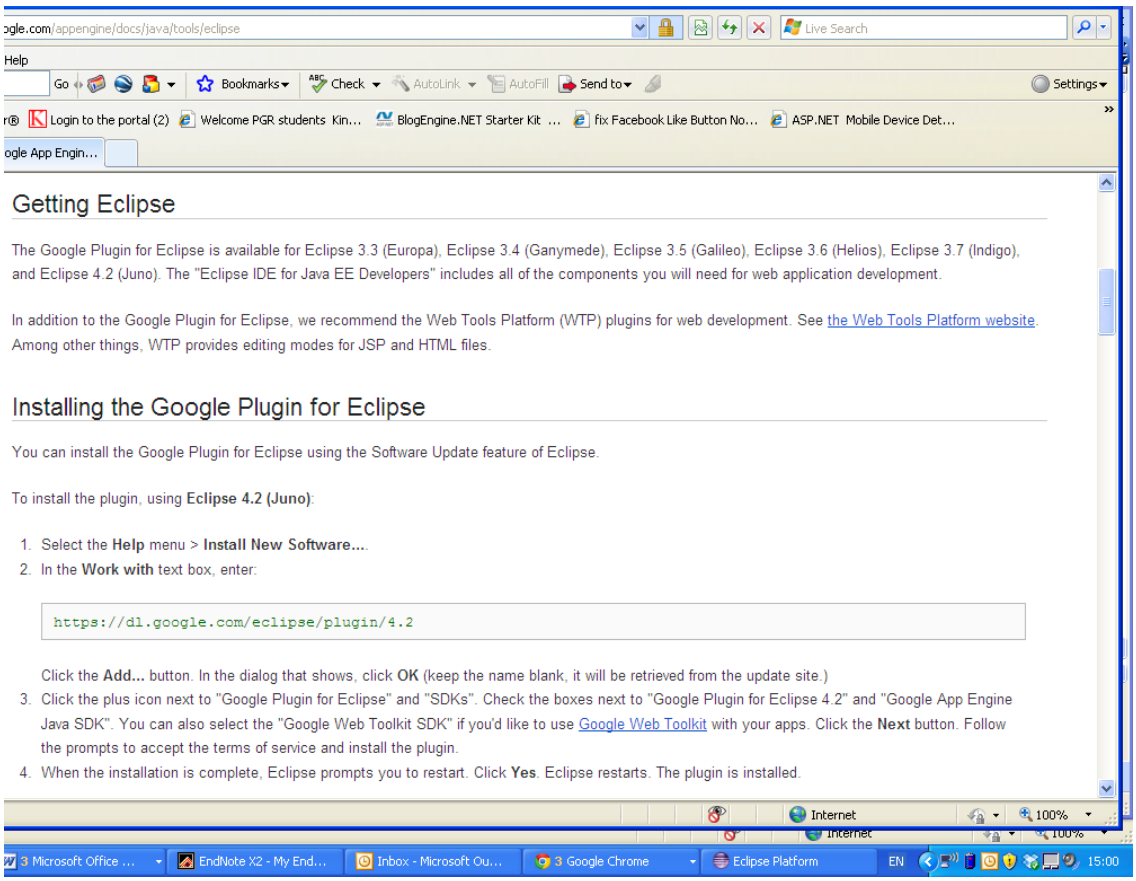
Package	Size	MD5 Checksum
ADT-20.0.3.zip	12390954 bytes	869a536b1c56d0cd920ed9ae259ae619

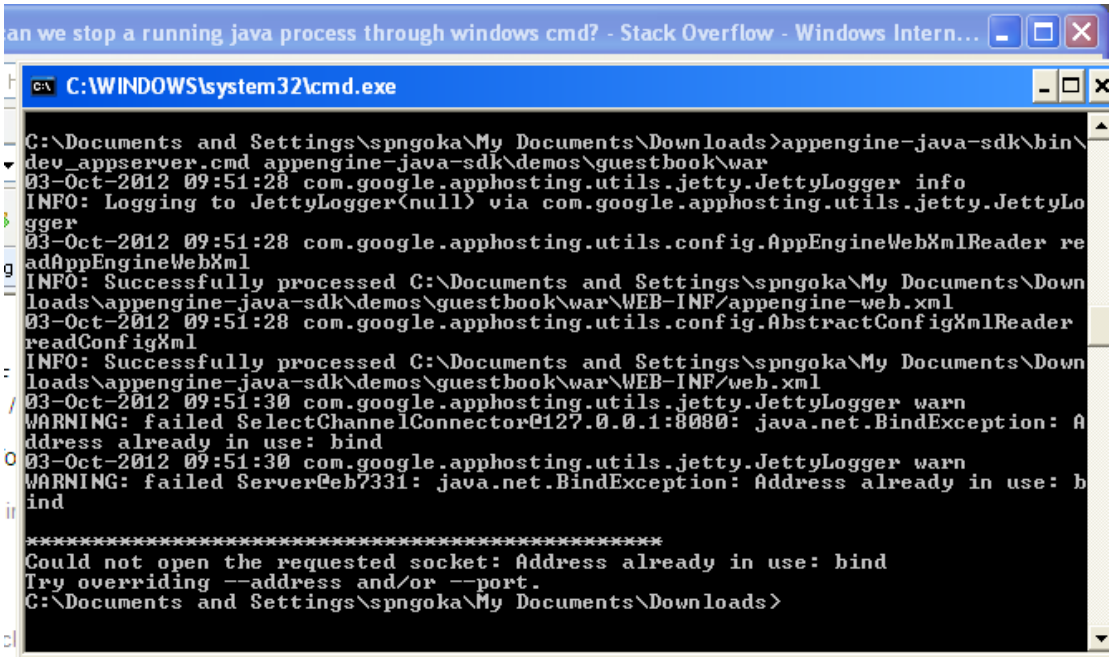
2. Start Eclipse, then select **Help** > **Install New Software**.
3. Click **Add**, in the top-right corner.
4. In the Add Repository dialog, click **Archive**.
5. Select the downloaded ADT-20.0.3.zip file and click **OK**.
6. Enter "ADT Plugin" for the name and click **OK**.
7. In the Available Software dialog, select the checkbox next to Developer Tools and click **Next**.
8. In the next window, you'll see a list of the tools to be downloaded. Click **Next**.
9. Read and accept the license agreements, then click **Finish**.
   If you get a security warning saying that the authenticity or validity of the software can't be established, click **OK**.
10. When the installation completes, restart Eclipse.

To update your plugin once you've installed using the zip file, you will have to follow these steps again instead of the default update instructions.

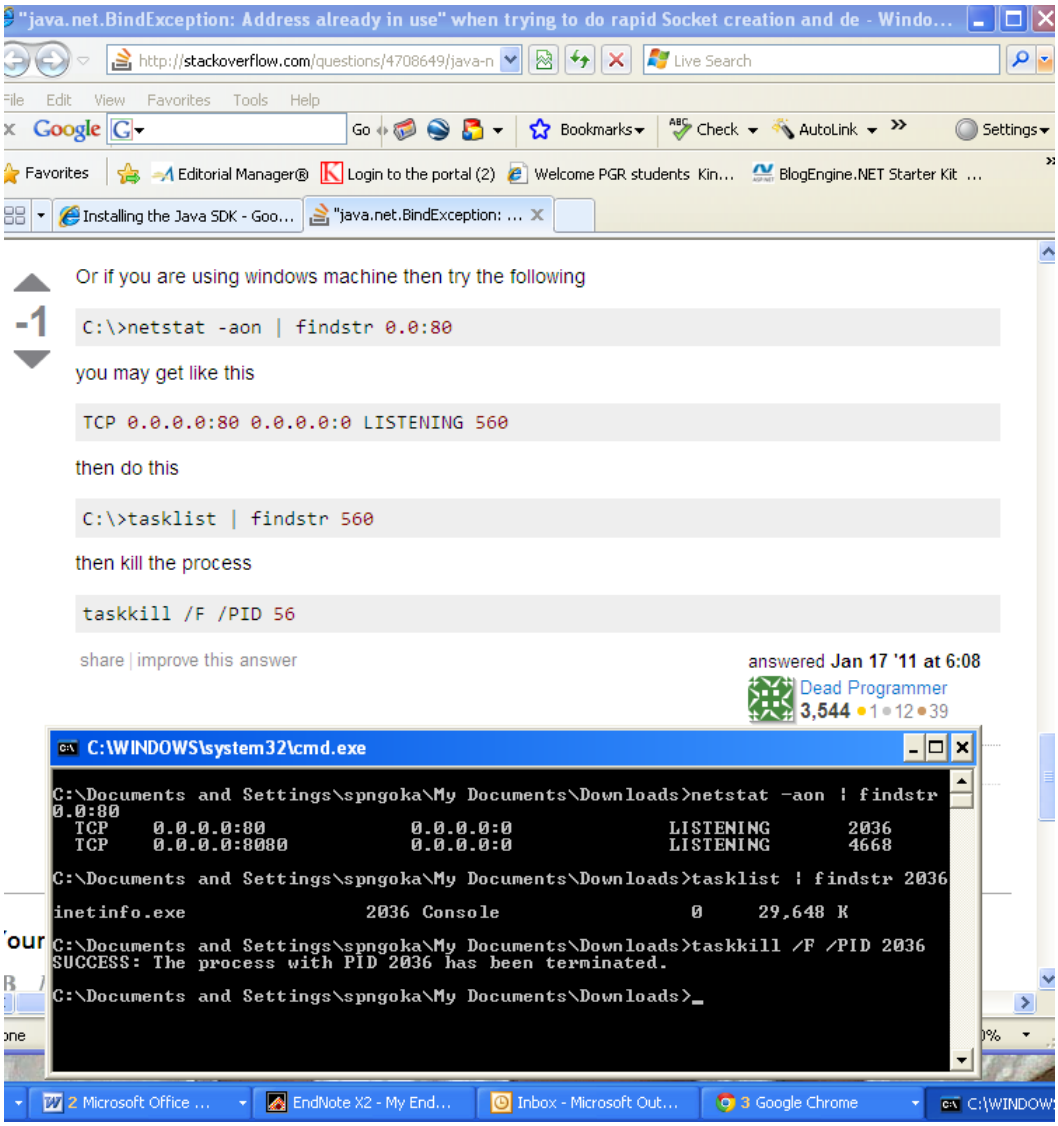
Using Google Engine for app development



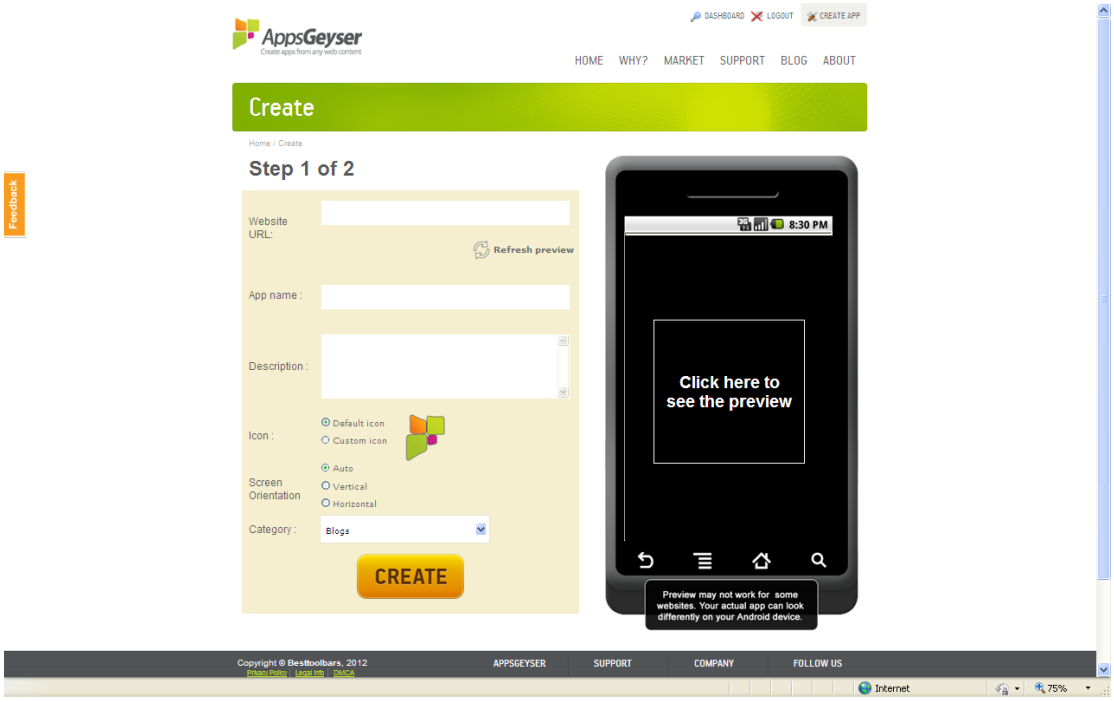


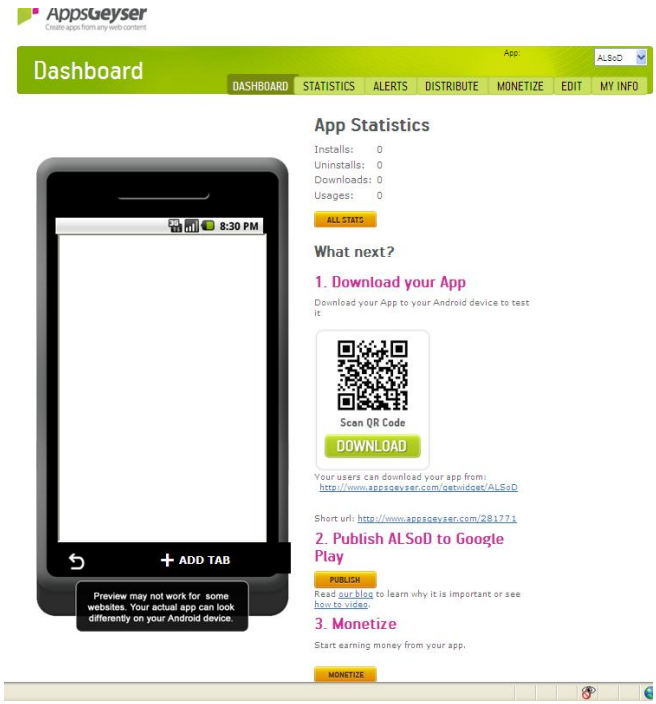


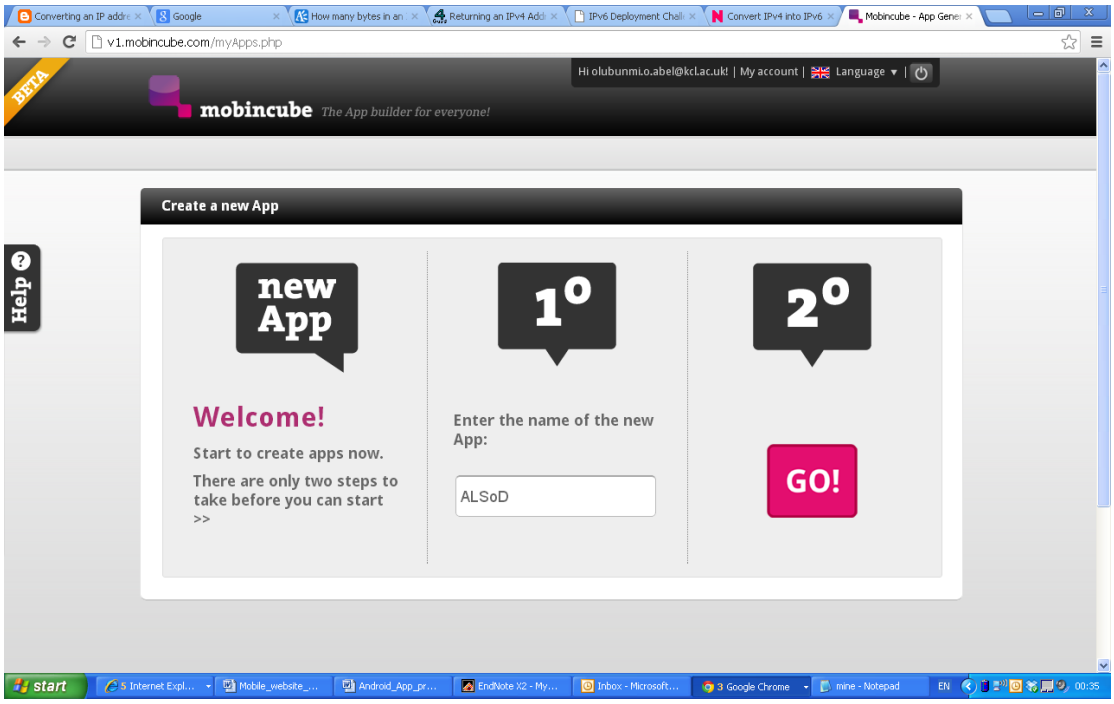
checked all java processes

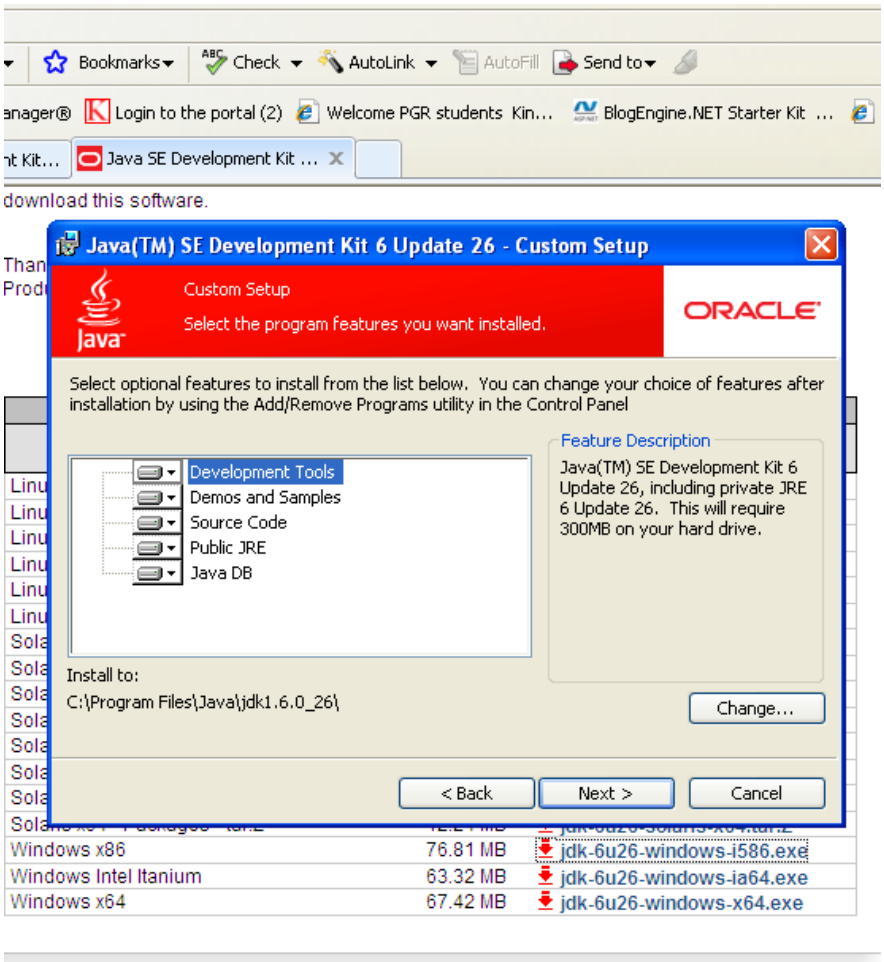

"java.net.BindException: Address already in use" when trying to do rapid Socket creation and de - Windo...
http://stackoverflow.com/questions/4708649/java-n
Live Search
File Edit View Favorites Tools Help
Google
Go
Bookmarks
Check
AutoLink
Settings
Favorites
Editorial Manager®
Login to the portal (2)
Welcome PGR students Kin...
BlogEngine.NET Starter Kit ...
Installing the Java SDK - Goo...
"java.net.BindException: ...
-1
Or if you are using windows machine then try the following
C:\>netstat -aon | findstr 0.0:80
you may get like this
TCP 0.0.0.0:80 0.0.0.0:0 LISTENING 560
then do this
C:\>tasklist | findstr 560
then kill the process
taskkill /F /PID 56
share | improve this answer
answered Jan 17 '11 at 6:08
Dead Programmer
3,544 1 12 39
C:\WINDOWS\system32\cmd.exe
C:\Documents and Settings\spngoka\My Documents\Downloads>netstat -aon | findstr 0.0:80
TCP 0.0.0.0:80 0.0.0.0:0 LISTENING 2036
TCP 0.0.0.0:8080 0.0.0.0:0 LISTENING 4668
C:\Documents and Settings\spngoka\My Documents\Downloads>tasklist | findstr 2036
inetinfo.exe 2036 Console 0 29,648 K
C:\Documents and Settings\spngoka\My Documents\Downloads>taskkill /F /PID 2036
SUCCESS: The process with PID 2036 has been terminated.
C:\Documents and Settings\spngoka\My Documents\Downloads>
2 Microsoft Office ...
EndNote X2 - My End...
Inbox - Microsoft Out...
3 Google Chrome
C:\WINDOW

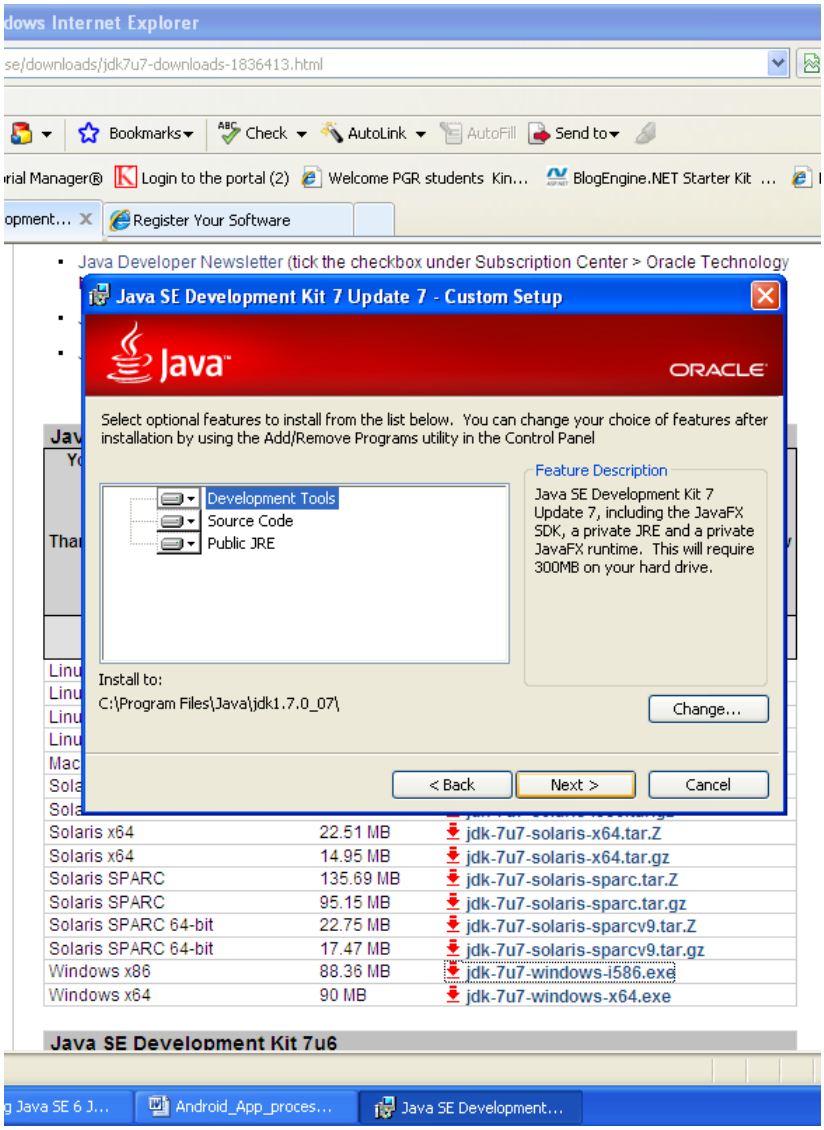

AppsGeyser
Create apps from any web content
DASHBOARD
LOGOUT
CREATE APP
HOME
WHY?
MARKET
SUPPORT
BLOG
ABOUT
Create App From
Home / Create App From
WEBSITE
ENTER URL
HTML CODE
CREATE OR COPY&PASTE HTML
YouTube Videos
ENTER KEYWORDS OR USE CHANNEL URL
more options...
Watch the Tour
How to create an App from Mobile Website
WATCH THE TOUR
How to create an App using HTML code or web widget
WATCH THE TOUR
Note: You can add more tabs for your app later
Feedback
Copyright © Besttoolbars, 2012
Privacy Policy | Legal Info | DMCA
APPSGEYSER
Why?
SUPPORT
FAQ
COMPANY
About
FOLLOW US
Blog
Done
Internet
100%

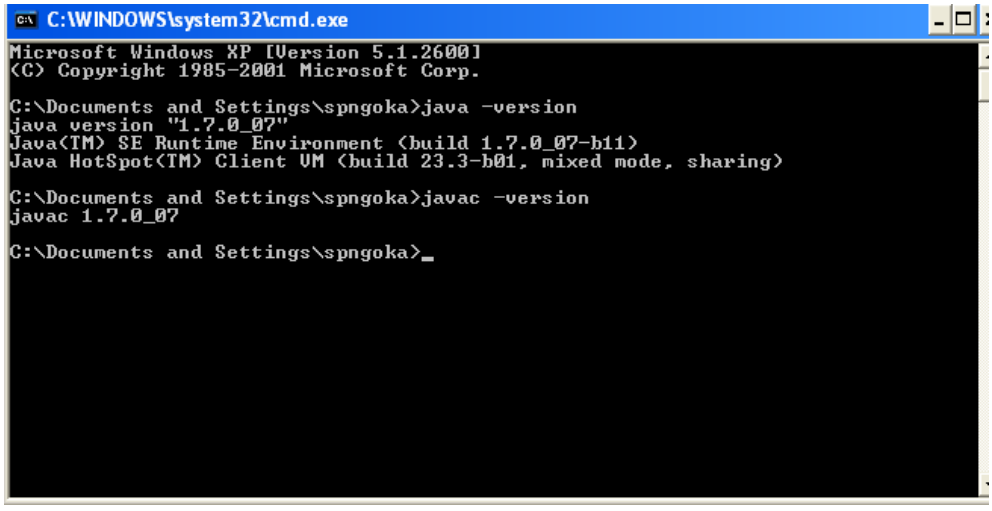

AppsGeyser
Create apps from any web content
DASHBOARD
LOGOUT
CREATE APP
HOME
WHY?
MARKET
SUPPORT
BLOG
ABOUT
Create
Home / Create
Step 1 of 2
Website URL:
Refresh preview
App name :
Description :
Icon :
Default icon
Custom icon
Screen Orientation
Auto
Vertical
Horizontal
Category :
Blogs
CREATE
8:30 PM
Click here to see the preview
Preview may not work for some websites. Your actual app can look differently on your Android device.
Feedback
Copyright © Besttoolbars, 2012
Privacy Policy | Legal Info | DMCA
APPSGEYSER
SUPPORT
COMPANY
FOLLOW US
Internet
75%

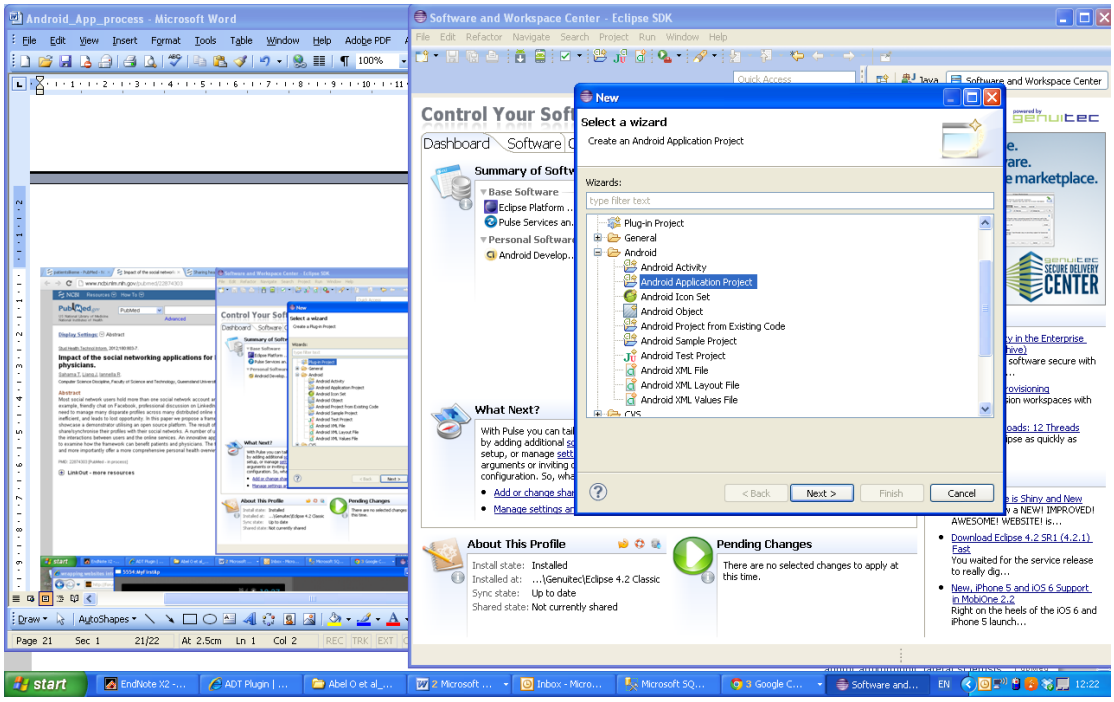

AppsGeyser
Dashboard
App:
ALSoD
DASHBOARD
STATISTICS
ALERTS
DISTRIBUTE
MONETIZE
EDIT
MY INFO
App Statistics
Installs: 0
Uninstalls: 0
Downloads: 0
Usages: 0
ALL STATS
What next?
1. Download your App
Download your App to your Android device to test it
Scan QR Code
DOWNLOAD
Your users can download your app from:
http://www.appsgeyser.com/getwidget/ALSoD
Short url: http://www.appsgeyser.com/281771
2. Publish ALSoD to Google Play
PUBLISH
Read our blog to learn why it is important or see how to video.
3. Monetize
Start earning money from your app.
MONETIZE
8:30 PM
ADD TAB
Preview may not work for some websites. Your actual app can look differently on your Android device.

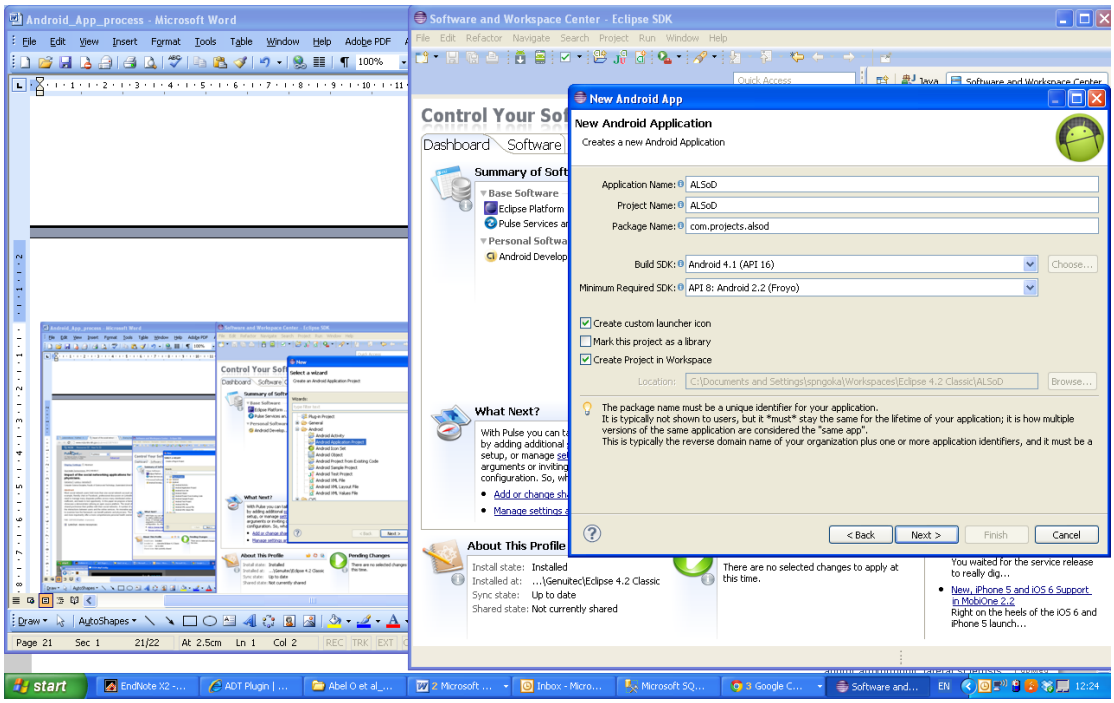

v1.mobincube.com/myApps.php
BETA
mobincube The App builder for everyone!
Hi olubunmi.o.abel@kcl.ac.uk! | My account | Language
Help
Create a new App
new App
1º
2º
Welcome!
Start to create apps now.
There are only two steps to take before you can start >>
Enter the name of the new App:
ALSoD
GO!
start

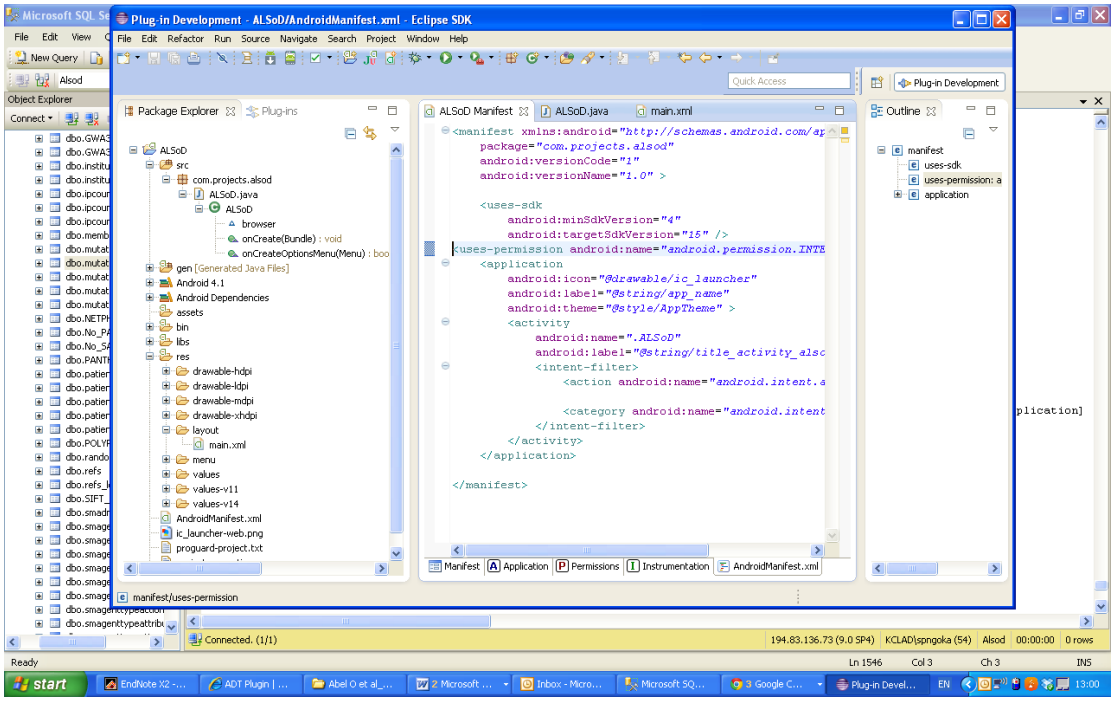

Bookmarks
Check
AutoLink
AutoFill
Send to
Login to the portal (2)
Welcome PGR students Kin...
BlogEngine.NET Starter Kit ...
Java SE Development Kit ...
download this software.
Java(TM) SE Development Kit 6 Update 26 - Custom Setup
Custom Setup
Select the program features you want installed.
ORACLE
Select optional features to install from the list below. You can change your choice of features after installation by using the Add/Remove Programs utility in the Control Panel
Development Tools
Demos and Samples
Source Code
Public JRE
Java DB
Feature Description
Java(TM) SE Development Kit 6 Update 26, including private JRE 6 Update 26. This will require 300MB on your hard drive.
Install to:
C:\Program Files\Java\jdk1.6.0_26\
Change...
< Back
Next >
Cancel
Windows x86 76.81 MB jdk-6u26-windows-i586.exe
Windows Intel Itanium 63.32 MB jdk-6u26-windows-ia64.exe
Windows x64 67.42 MB jdk-6u26-windows-x64.exe

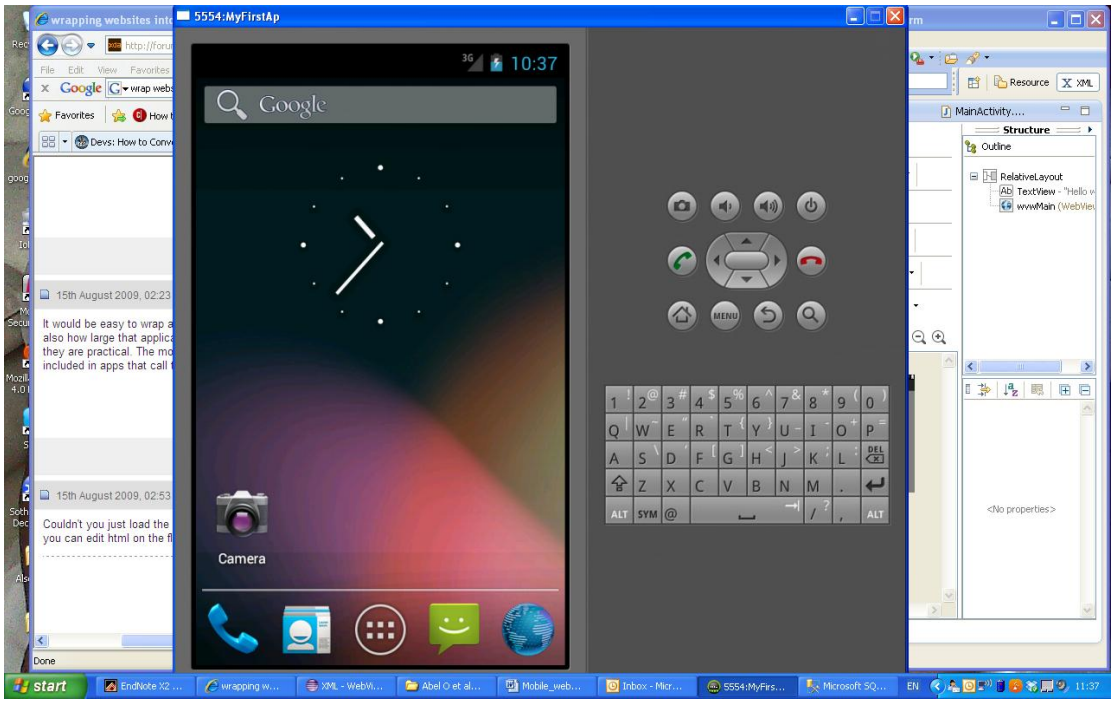

dows Internet Explorer
se/downloads/jdk7u7-downloads-1836413.html
Bookmarks
Check
AutoLink
AutoFill
Send to
rial Manager®
Login to the portal (2)
Welcome PGR students Kin...
BlogEngine.NET Starter Kit ...
opment...
Register Your Software
Java Developer Newsletter (tick the checkbox under Subscription Center > Oracle Technology
Java SE Development Kit 7 Update 7 - Custom Setup
Java
ORACLE
Select optional features to install from the list below. You can change your choice of features after installation by using the Add/Remove Programs utility in the Control Panel
Development Tools
Source Code
Public JRE
Feature Description
Java SE Development Kit 7 Update 7, including the JavaFX SDK, a private JRE and a private JavaFX runtime. This will require 300MB on your hard drive.
Install to:
C:\Program Files\Java\jdk1.7.0_07\
Change...
< Back
Next >
Cancel
Solaris x64 22.51 MB jdk-7u7-solaris-x64.tar.Z
Solaris x64 14.95 MB jdk-7u7-solaris-x64.tar.gz
Solaris SPARC 135.69 MB jdk-7u7-solaris-sparc.tar.Z
Solaris SPARC 95.15 MB jdk-7u7-solaris-sparc.tar.gz
Solaris SPARC 64-bit 22.75 MB jdk-7u7-solaris-sparcv9.tar.Z
Solaris SPARC 64-bit 17.47 MB jdk-7u7-solaris-sparcv9.tar.gz
Windows x86 88.36 MB jdk-7u7-windows-i586.exe
Windows x64 90 MB jdk-7u7-windows-x64.exe
Java SE Development Kit 7u6
Java SE 6 J...
Android_App_proces...
Java SE Development...

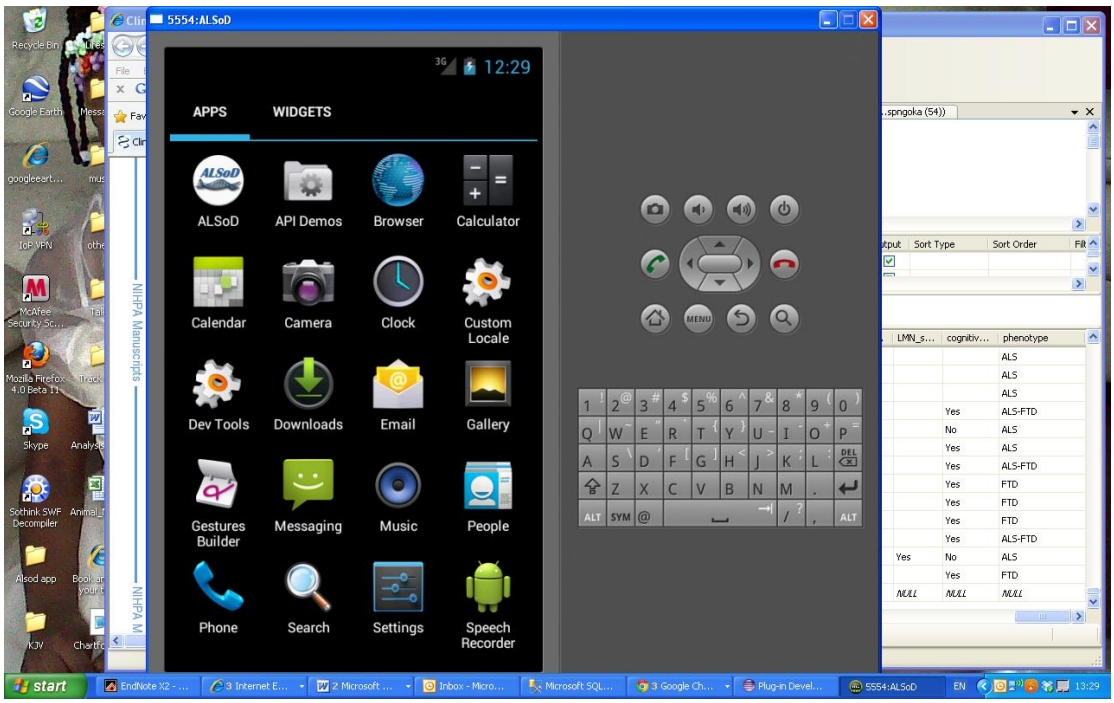

C:\WINDOWS\system32\cmd.exe
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.
C:\Documents and Settings\spngoka>java -version
java version "1.7.0_07"
Java(TM) SE Runtime Environment (build 1.7.0_07-b11)
Java HotSpot(TM) Client VM (build 23.3-b01, mixed mode, sharing)
C:\Documents and Settings\spngoka>javac -version
javac 1.7.0_07
C:\Documents and Settings\spngoka>

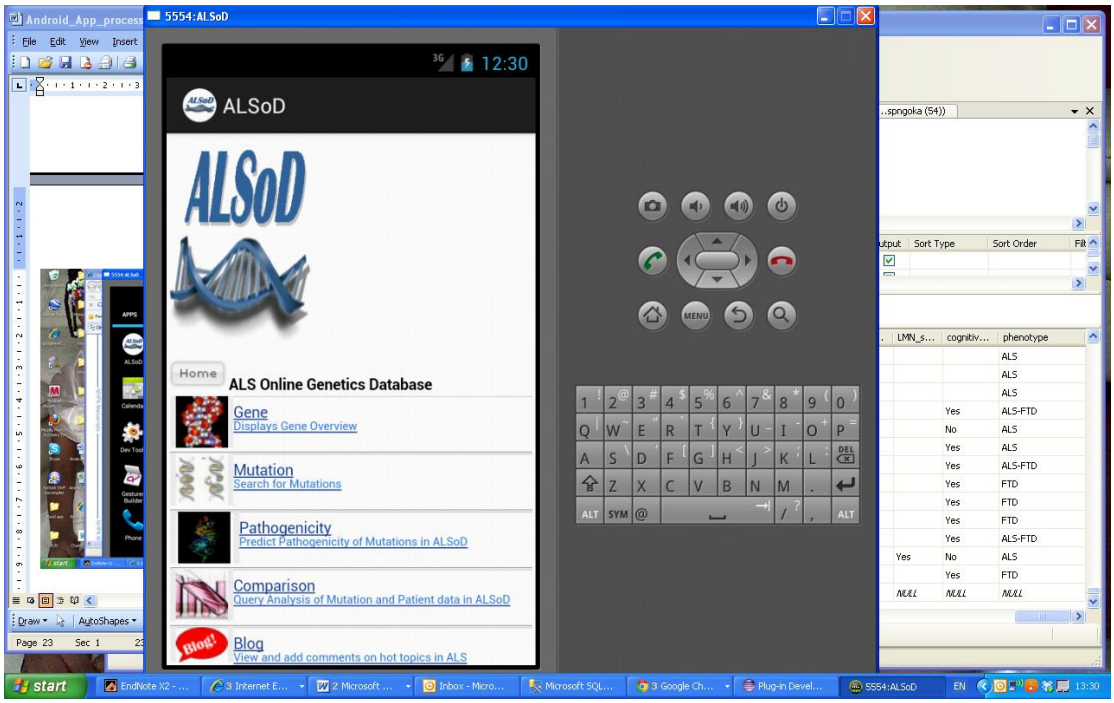

Android_App_process - Microsoft Word
Software and Workspace Center - Eclipse SDK
New
Select a wizard
Create an Android Application Project
Wizards:
type filter text
Plug-in Project
General
Android
Android Activity
Android Application Project
Android Icon Set
Android Object
Android Project from Existing Code
Android Sample Project
Android Test Project
Android XML File
Android XML Layout File
Android XML Values File
< Back
Next >
Finish
Cancel
Control Your Soft
Dashboard
Software
Summary of Softw
Base Software
Eclipse Platform
Pulse Services an
Personal Software
Android Develop
What Next?
With Pulse you can
by adding additional
setup, or manage
arguments or inviting
configuration. So, wh
Add or change sha
Manage settings a
About This Profile
Install state: Installed
Installed at: ...\Genuitec\Eclipse 4.2 Classic
Sync state: Up to date
Shared state: Not currently shared
Pending Changes
There are no selected changes to apply at this time.
Page 21 Sec 1 21/22 At 2.5cm Ln 1 Col 2
start
12:22


Android_App_process - Microsoft Word
Software and Workspace Center - Eclipse SDK
New Android App
New Android Application
Creates a new Android Application
Application Name: ALSoD
Project Name: ALSoD
Package Name: com.projects.alsod
Build SDK: Android 4.1 (API 16)
Minimum Required SDK: API 8: Android 2.2 (Froyo)
Create custom launcher icon
Mark this project as a library
Create Project in Workspace
Location: C:\Documents and Settings\spngoka\Workspaces\Eclipse 4.2 Classic\ALSoD
Browse...
The package name must be a unique identifier for your application.
It is typically not shown to users, but it *must* stay the same for the lifetime of your application; it is how multiple versions of the same application are considered the "same app".
This is typically the reverse domain name of your organization plus one or more application identifiers, and it must be a
< Back
Next >
Finish
Cancel
Control Your Soft
Dashboard
Software
Summary of Soft
Base Software
Eclipse Platform
Pulse Services ar
Personal Softwa
Android Develop
What Next?
With Pulse you can t
by adding additional
setup, or manage se
arguments or invitin
configuration. So, wh
Add or change sha
Manage settings a
About This Profile
Install state: Installed
Installed at: ...\Genuitec\Eclipse 4.2 Classic
Sync state: Up to date
Shared state: Not currently shared
There are no selected changes to apply at this time.
Page 21 Sec 1 21/22 At 2.5cm Ln 1 Col 2
start
12:24

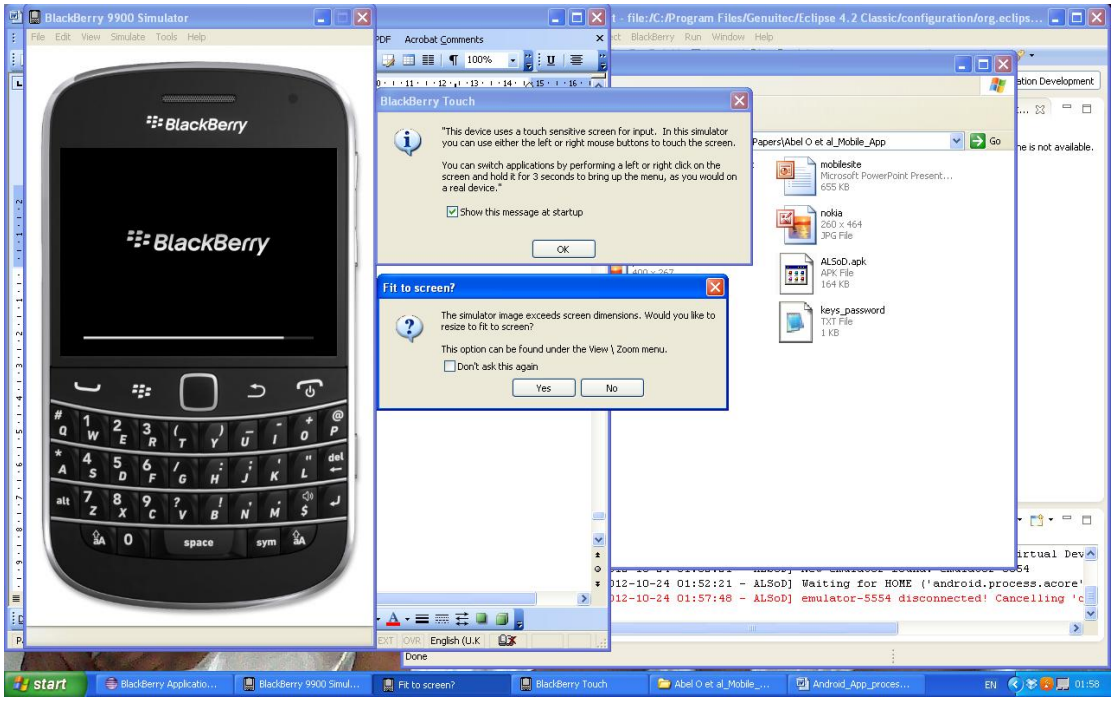

Plug-in Development - ALSoD/AndroidManifest.xml - Eclipse SDK
File Edit Refactor Run Source Navigate Search Project Window Help
Package Explorer
ALSoD Manifest
ALSoD.java
main.xml
Outline
Manifest
Application
Permissions
Instrumentation
AndroidManifest.xml

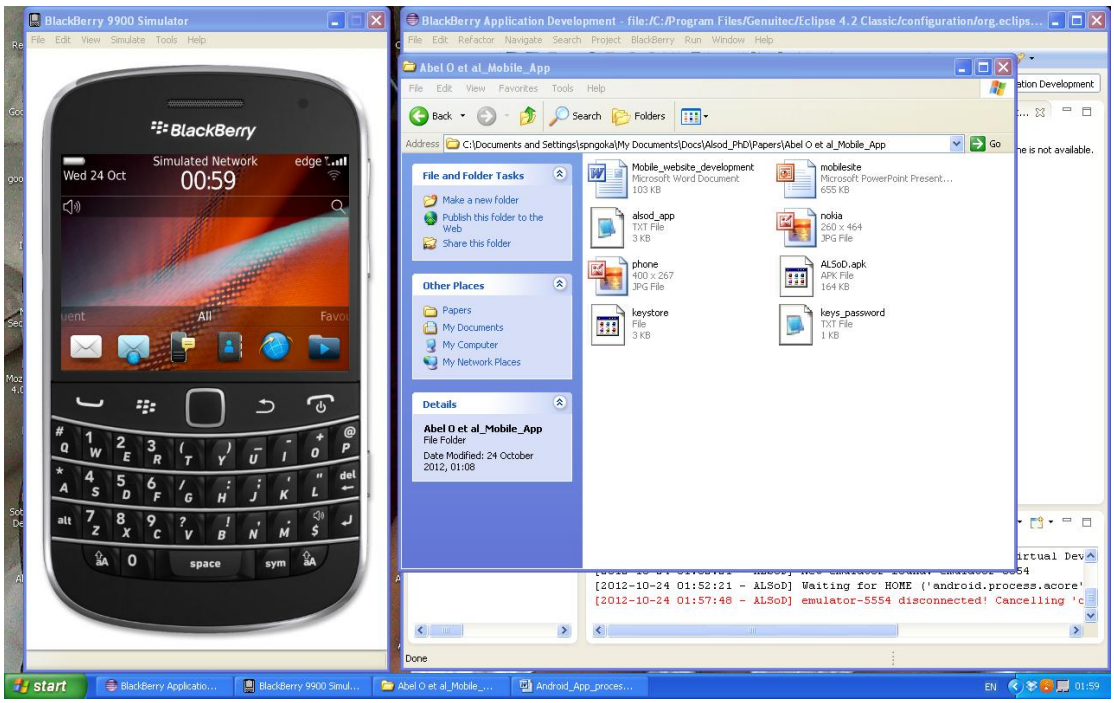

5554:MyFirstAp
10:37
Google
Camera
MainActivity....
Structure
Outline
RelativeLayout
TextView
wvwMain
<No properties>


5554:ALSoD
12:29
APPS
WIDGETS
ALSoD
API Demos
Browser
Calculator
Calendar
Camera
Clock
Custom Locale
Dev Tools
Downloads
Email
Gallery
Gestures Builder
Messaging
Music
People
Phone
Search
Settings
Speech Recorder


5554:ALSoD
12:30
ALSoD
Home
ALS Online Genetics Database
Gene
Displays Gene Overview
Mutation
Search for Mutations
Pathogenicity
Predict Pathogenicity of Mutations in ALSoD
Comparison
Query Analysis of Mutation and Patient data in ALSoD
Blog
View and add comments on hot topics in ALS


BlackBerry 9900 Simulator
File Edit View Simulate Tools Help
BlackBerry
BlackBerry
Click and drag to rotate or tilt


BlackBerry 9900 Simulator
BlackBerry Touch
"This device uses a touch sensitive screen for input. In this simulator you can use either the left or right mouse buttons to touch the screen.
You can switch applications by performing a left or right click on the screen and hold it for 3 seconds to bring up the menu, as you would on a real device."
Show this message at startup
OK
Fit to screen?
The simulator image exceeds screen dimensions. Would you like to resize to fit to screen?
This option can be found under the View \ Zoom menu.
Don't ask this again
Yes
No
mobilesite
nokia
ALSoD.apk
keys_password
start


BlackBerry 9900 Simulator
BlackBerry Application Development - file:/C:/Program Files/Genuitec/Eclipse 4.2 Classic/configuration/org.eclips...
Abel O et al_Mobile_App
Simulated Network
Wed 24 Oct
00:59
File and Folder Tasks
Other Places
Details
Mobile_website_development
mobilesite
alsod_app
nokia
phone
ALSoD.apk
keystore
keys_password
[2012-10-24 01:52:21 - ALSoD] Waiting for HOME ('android.process.acore'
[2012-10-24 01:57:48 - ALSoD] emulator-5554 disconnected! Cancelling 'c
Done
start

Appendix 11 – SurveyMonkey screenshot

***1. Please score each of these genes according to how credible they are as established ALS genes for typical ALS, with 1 being most credible, and 14 being least credible. You may score more than one gene with the same score.**

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
VAPB	○	○	○	○	○	○	○	○	○	○	○	○	○	○
SOD1	○	○	○	○	○	○	○	○	○	○	○	○	○	○
DAO	○	○	○	○	○	○	○	○	○	○	○	○	○	○
OPTN	○	○	○	○	○	○	○	○	○	○	○	○	○	○
SETX	○	○	○	○	○	○	○	○	○	○	○	○	○	○
TARDBP(TDP43)	○	○	○	○	○	○	○	○	○	○	○	○	○	○
NEFH	○	○	○	○	○	○	○	○	○	○	○	○	○	○
VCP	○	○	○	○	○	○	○	○	○	○	○	○	○	○
ALS2	○	○	○	○	○	○	○	○	○	○	○	○	○	○
FUS	○	○	○	○	○	○	○	○	○	○	○	○	○	○
SPG11	○	○	○	○	○	○	○	○	○	○	○	○	○	○
ANG	○	○	○	○	○	○	○	○	○	○	○	○	○	○
DCTN1	○	○	○	○	○	○	○	○	○	○	○	○	○	○
FIG4	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Done

Appendix 12 – Script for Blog

```
<script language="c#" runat="server">
protected void Page_Load(object sender, System.EventArgs e)
        {
            // Put user code to initialize the page here
            if (IsPostBack)
            {

            }
            else
            {
                // initialize Blog and read it into a dataset
                DataSet ds = ReadBlogIntoTable();

                // if we created a new entry, append it to the BlogList
                // and save it to the xml file
                if ( (bool)Session["Changed"])
                {
                    AppendComments(ds);
                    WriteXmlComments(ds);
                    Session["Changed"] = false;
                }

                // Dynamically build the blog into a WebControls.Table
                RebuildTableView(ds);
            }
        }

        DataSet ReadBlogIntoTable()
        {
            return ReadXmlComments();
        }
```

```
private DataSet ReadXmlComments()
{
    // construct the DataSet
    DataSet ds = new DataSet();

    // form the server path of the feedback database
    string filename = Server.MapPath(".\\") + "comments.xml";

    // if the file exists, read  the blog database into the data set
    if (File.Exists(filename))
    {
        ds.ReadXml(filename);
    }
    else
    {
        // otherwise we need to create a new new in-memory
        // Database table from scratch
        DataTable theTable = new DataTable("Comments");
        ds.Tables.Add(theTable);

        // add a name, time, title, and blog columns to our mock-up blog
database

        theTable.Columns.Add("Name", Type.GetType("System.String"));
        theTable.Columns.Add("Time", Type.GetType("System.DateTime"));
        theTable.Columns.Add("Title", Type.GetType("System.String"));
        theTable.Columns.Add("Blog", Type.GetType("System.String"));
    }

    return ds;
}

private void WriteXmlComments(DataSet ds)
{
    string filename = Server.MapPath(".\\") + "comments.xml";
```

```
            ds.WriteXml(filename);
}

/// <summary>
///  Append comments to xml file
/// </summary>
/// <param name="ds"></param>
void AppendComments(DataSet ds)
{
            // create a new DataRow
            DataRow dr = ds.Tables["Comments"].NewRow();
            // Populate the row from the text boxes filled by the user
            dr[0] = Session["Name"];
            dr[1] = DateTime.Now;
            dr[2] = Session["Title"];
            dr[3] = Session["Blog"];

            // add the row to the DataSet
            ds.Tables["Comments"].Rows.InsertAt(dr, 0);

            WriteXmlComments(ds);
}

void RebuildTableView(DataSet ds)
{
            string previousUser = "";  // track previous

            // loop through each row in the data set and create
            // the row on the web page in a web table
            foreach (DataRow dr in ds.Tables[0].Rows)
            {
                        // add title (use a single column)
                        TableRow tr = new TableRow();
```

```
                tr.Cells.Add(new TableCell());

                // change title color slightly
                tr.Cells[0].ForeColor = Color.Navy;
                tr.Cells[0].Width = 500;

            //  make the text title big and purple
                tr.Cells[0].Text = "<FONT SIZE=5 COLOR=purple FACE=Rockwell><B>" + dr[2].ToString() + "</B></FONT>";
                this.BlogTable.Rows.Add(tr);

                // add blog in a single column and span 2 columns
                tr = new TableRow();
                tr.Cells.Add(new TableCell());
                tr.Cells[0].Width = 550;
                tr.Cells[0].ColumnSpan = 2;
                tr.Cells[0].Text = dr[3].ToString();
                this.BlogTable.Rows.Add(tr);

                // add user who posted and date (use two columns in the row)
                tr = new TableRow();
                tr.Height = 50;
                tr.HorizontalAlign = HorizontalAlign.Left;
                tr.VerticalAlign = VerticalAlign.Bottom;
                tr.Cells.Add(new TableCell());
                tr.Cells.Add(new TableCell());

                tr.Cells[0].Text = "Posted by " + dr[0].ToString();
                DateTime postTime = DateTime.Parse(dr[1].ToString());
                tr.Cells[1].HorizontalAlign = HorizontalAlign.Right;

                tr.Cells[1].Text = String.Format("<i>{0}</i>", postTime.ToString("MMM dd, 2005 @ hh:mm"));
```

```
                this.BlogTable.Rows.Add(tr);

                // add separtor graphic and span 2 columns
                tr = new TableRow();
                tr.Cells.Add(new TableCell());
                tr.Cells.Add(new TableCell());
                tr.Cells[0].ColumnSpan  = 2;

                this.BlogTable.Rows.Add(tr);

                string imageFile = this.ResolveUrl(
@".\images\separator.jpg");
                System.Web.UI.WebControls.Image separator = new
System.Web.UI.WebControls.Image();
                separator.ImageUrl = imageFile;
                separator.Width = 600;
                separator.Height = 32;
                separator.Visible = true;

                tr.Cells[0].Controls.Add(separator);
                tr.Cells[0].HorizontalAlign = HorizontalAlign.Center;


            }

        }


        #region Web Form Designer generated code
        override protected void OnInit(EventArgs e)
        {
            //
            // CODEGEN: This call is required by the ASP.NET Web Form Designer.
            //
```

```
        InitializeComponent();
        base.OnInit(e);
    }

    /// <summary>
    /// Required method for Designer support - do not modify
    /// the contents of this method with the code editor.
    /// </summary>
    private void InitializeComponent()
    {

    }
    #endregion

    protected void btnOpenEntry_Click(object sender, System.EventArgs e)
    {
        Response.Redirect("BlogEntry.aspx");
    }

</script>
```

Appendix 13 – Masterpage Websitemap

Treeview

```
<%@ Page Language="C#" MasterPageFile="~/master/MasterPage.master"%>

<asp:Content ID="Content1" ContentPlaceHolderID="ContentPlaceHolder1"
Runat="Server">
<asp:TreeView ID="TreeView1" runat="server">
          <Nodes>
            <asp:TreeNode Text="ALSoD Masterpage">
            <asp:TreeNode Text="Homepage"
NavigateUrl="http://alsod.iop.kcl.ac.uk/index.aspx"/>
              <asp:TreeNode Text="Analysis">
                <asp:TreeNode Text="Detailed Analysis"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/analysis.aspx" />
                <asp:TreeNode Text="Interactions"
NavigateUrl="http://alsod.iop.kcl.ac.uk/overview/interaction.aspx"/>
                <asp:TreeNode Text="Side-by-side comparison"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/statistics.aspx" />
                <asp:TreeNode Text="Predict Pathogenicity"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/pathogenicity.aspx" />
                <asp:TreeNode Text="Gene Credibility"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/credibility.aspx" />
                </asp:TreeNode>
               <asp:TreeNode Text="Summary">
                <asp:TreeNode Text="Genetic Data Report"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/report.aspx" />
                <asp:TreeNode Text="Search"
NavigateUrl="http://alsod.iop.kcl.ac.uk/index7.aspx"/>
                <asp:TreeNode Text="Database Summary"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Summary/summary.aspx" />
                </asp:TreeNode>
                <asp:TreeNode Text="GWAS">
                <asp:TreeNode Text="GWAS 5 populations"
NavigateUrl="http://alsod.iop.kcl.ac.uk/GWA2/index.aspx" />
                <asp:TreeNode Text="GWAS by Fogh 2013"
NavigateUrl="http://alsod.iop.kcl.ac.uk/GWA2/gwas_fogh.aspx"/>
                <asp:TreeNode Text="Help"
NavigateUrl="http://alsod.iop.kcl.ac.uk/misc/analysiserror.aspx" />
                </asp:TreeNode>
                <asp:TreeNode Text="News">
                <asp:TreeNode Text="Latest News"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Search/searchWeb.aspx" >
                    <asp:TreeNode Text="alsforum.com"
NavigateUrl="http://www.alsforums.com/forum/als-research-news/8373-first-
neuroprotective-gene-patients-amyotrophic-lateral-sclerosis-isolated.html
"/>
                    <asp:TreeNode Text="als.net forum"
NavigateUrl="http://www.als.net/forum/ "/>
                    <asp:TreeNode Text="bbc.co.uk"
NavigateUrl="http://news.bbc.co.uk/1/hi/health/8045795.stm"/>
                    <asp:TreeNode Text="medical news"
NavigateUrl="http://alsod.iop.kcl.ac.uk/misc/diseaseDetails.aspx"/>
                    <asp:TreeNode Text="RSS feeds"
NavigateUrl="http://pipes.yahoo.com/pipes/pipe.run?_id=7d74ac65a4586449b1d7
579dfbae9beb&_render=rss" />
                    <asp:TreeNode Text="Google Search" >
                        <asp:TreeNode Text="News" />
                        <asp:TreeNode Text="Web" />
                        <asp:TreeNode Text="Blog" />
```

```
                        <asp:TreeNode Text="Book" />
                        <asp:TreeNode Text="Patent" />
                    </asp:TreeNode>
                </asp:TreeNode>
                <asp:TreeNode Text="Blogs"
NavigateUrl="http://alsod.iop.kcl.ac.uk/BlogList.aspx" />
                <asp:TreeNode Text="Social Media"
NavigateUrl="http://alsod.iop.kcl.ac.uk/misc/socialmedia.aspx" >
                    <asp:TreeNode Text="Facebook" />
                    <asp:TreeNode Text="Twitter" />
                </asp:TreeNode>
                </asp:TreeNode>
                <asp:TreeNode Text="Data">
                <asp:TreeNode Text="Detailed Analysis"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/analysis.aspx" />
                <asp:TreeNode Text="Interactions"
NavigateUrl="http://alsod.iop.kcl.ac.uk/overview/interaction.aspx"/>
                <asp:TreeNode Text="Side-by-side comparison"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/statistics.aspx" />
                <asp:TreeNode Text="Predict Pathogenicity"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/pathogenicity.aspx" />
                <asp:TreeNode Text="Gene Credibility"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/credibility.aspx" />
                </asp:TreeNode>
                <asp:TreeNode Text="Contributors">
                <asp:TreeNode Text="Detailed Analysis"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/analysis.aspx" />
                <asp:TreeNode Text="Interactions"
NavigateUrl="http://alsod.iop.kcl.ac.uk/overview/interaction.aspx"/>
                <asp:TreeNode Text="Side-by-side comparison"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/statistics.aspx" />
                <asp:TreeNode Text="Predict Pathogenicity"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/pathogenicity.aspx" />
                <asp:TreeNode Text="Gene Credibility"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/credibility.aspx" />
                </asp:TreeNode>
                <asp:TreeNode Text="Feedback">
                <asp:TreeNode Text="Detailed Analysis"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/analysis.aspx" />
                <asp:TreeNode Text="Interactions"
NavigateUrl="http://alsod.iop.kcl.ac.uk/overview/interaction.aspx"/>
                <asp:TreeNode Text="Side-by-side comparison"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/statistics.aspx" />
                <asp:TreeNode Text="Predict Pathogenicity"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/pathogenicity.aspx" />
                <asp:TreeNode Text="Gene Credibility"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Statistics/credibility.aspx" />
                </asp:TreeNode>
            </asp:TreeNode>
            <asp:TreeNode Text="Top section">
                <asp:TreeNode Text="Login"
NavigateUrl="http://alsod.iop.kcl.ac.uk/login/loginAuthenticate.aspx"/>
                <asp:TreeNode Text="Feedback"
NavigateUrl="http://alsod.iop.kcl.ac.uk/contact/guestbook.aspx"/>
                <asp:TreeNode Text="Help"
NavigateUrl="http://alsod.iop.kcl.ac.uk/misc/FAQs.aspx"/>
                <asp:TreeNode Text="Register"
NavigateUrl="http://alsod.iop.kcl.ac.uk/login/preSign0.aspx" />
                <asp:TreeNode Text="Cookies"
NavigateUrl="http://www.aboutcookies.org/default.aspx?page=5"/>
                <asp:TreeNode Text="Print" />
```

```
                <asp:TreeNode Text="Cite us"
NavigateUrl="http://alsod.iop.kcl.ac.uk/contact/contact.aspx"/>
                <asp:TreeNode Text="Sitemap"
NavigateUrl="http://alsod.iop.kcl.ac.uk/maps/treeview.aspx"/>
              </asp:TreeNode>
              <asp:TreeNode Text="Left section">
                <asp:TreeNode Text="Chromosomes">
                <asp:TreeNode Text="Chromosome 1"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo1.aspx"/>
                <asp:TreeNode Text="Chromosome 2"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo2.aspx"/>
                <asp:TreeNode Text="Chromosome 3"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo3.aspx"/>
                <asp:TreeNode Text="Chromosome 4"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo4.aspx"/>
                <asp:TreeNode Text="Chromosome 5"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo5.aspx"/>
                <asp:TreeNode Text="Chromosome 6"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo6.aspx"/>
                <asp:TreeNode Text="Chromosome 7"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo7.aspx"/>
                <asp:TreeNode Text="Chromosome 8"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo8.aspx"/>
                <asp:TreeNode Text="Chromosome 9"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo9.aspx"/>
                <asp:TreeNode Text="Chromosome 10"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo10.aspx"/>
                <asp:TreeNode Text="Chromosome 11"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo11.aspx"/>
                <asp:TreeNode Text="Chromosome 12"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo12.aspx"/>
                <asp:TreeNode Text="Chromosome 13"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo13.aspx"/>
                <asp:TreeNode Text="Chromosome 14"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo14.aspx"/>
                <asp:TreeNode Text="Chromosome 15"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo15.aspx"/>
                <asp:TreeNode Text="Chromosome 16"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo16.aspx"/>
                <asp:TreeNode Text="Chromosome 17"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo17.aspx"/>
                <asp:TreeNode Text="Chromosome 18"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo18.aspx"/>
                <asp:TreeNode Text="Chromosome 19"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo19.aspx"/>
                <asp:TreeNode Text="Chromosome 20"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo20.aspx"/>
                <asp:TreeNode Text="Chromosome 21"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo21.aspx"/>
                <asp:TreeNode Text="Chromosome 22"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromo22.aspx"/>
                <asp:TreeNode Text="Chromosome ALL"
NavigateUrl="http://alsod.iop.kcl.ac.uk/Chromosomes/chromoALL.aspx"/>
                </asp:TreeNode>
                <asp:TreeNode Text="Funders" />
                <asp:TreeNode Text="Visit Statistics" />
                <asp:TreeNode Text="Facebook" />
              </asp:TreeNode>
          </Nodes>
        </asp:TreeView>
</asp:Content>
```

Map on http://alsod.iop.kcl.ac.uk/maps/treeview.aspx

Login Feedback Help Register Sitemap Cookies Print Cite us

ALS ONLINE GENETICS DATABASE

You are here ==>

search Google ...type to search Google Click button to search ALSoD

NEW!!! Check ALSoD on your smartphones and tablets ..... ALSoD is now Mobile-Friendly!!! ..... as usual use http://alsod.iop.

Recommend Share 41

Home Analysis Summary GWAS News Data Contributors Feedback

Chromosomes

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y ALL

ALSoD is a joint project of World Federation of Neurology and European Network to Cure ALS. The work leading to these results has received funding from the European Community's Health Seventh Framework Programme FP7/2007-2013 under grant agreement number 259867.

ALSoD is funded by

Since 03/01/09
Viewed 925759 times,
Visited by
44637
Unique Users
You are accessing from
EUROPEAN UNION
Last Update
20th February 2014

41

Recommend

Share

- ALSoD Masterpage
  - Homepage
  - Analysis
    - Detailed Analysis
    - Interactions
    - Side-by-side comparison
    - Predict Pathogenicity
    - Gene Credibility
  - Summary
    - Genetic Data Report
    - Search
    - Database Summary
  - GWAS
    - GWAS 5 populations
    - GWAS by Fogh 2013
    - Help
  - News
    - Latest News
      - alsforum.com
      - als.net forum
      - bbc.co.uk
      - medical news
      - RSS feeds
      - Google Search
        - News
        - Web
        - Blog
        - Book
        - Patent
    - Blogs
    - Social Media
      - Facebook
      - Twitter
  - Data
    - Detailed Analysis
    - Interactions
    - Side-by-side comparison
    - Predict Pathogenicity
    - Gene Credibility
  - Contributors
    - Detailed Analysis
    - Interactions
    - Side-by-side comparison
    - Predict Pathogenicity
    - Gene Credibility
  - Feedback
    - Detailed Analysis
    - Interactions
    - Side-by-side comparison
    - Predict Pathogenicity
    - Gene Credibility
- Top section
  - Login
  - Feedback
  - Help
  - Register
  - Cookies
  - Print
  - Cite us
  - Sitemap
- Left section
  - Chromosomes
  - Fundars
  - Visit Statistics
  - Facebook

Map on http://alsod.iop.kcl.ac.uk/Web.sitemap

```
<?xml version="1.0" encoding="utf-8" ?>
<siteMap xmlns="http://schemas.microsoft.com/AspNet/SiteMap-File-1.0" >
    <siteMapNode url="~/index.aspx" title="Home"  description="">
        <siteMapNode url="~/reports/reportSummary.aspx" title="Genes"
description="" >
      <siteMapNode url="~/Applets/TARDBP(TDP43)species.aspx"
title="TARDBP(TDP43) Alignment and Mutation"  description="" />
      <siteMapNode url="~/Applets/SOD1species.aspx" title="SOD1 Alignment
and Mutation"  description="" />
      <siteMapNode url="~/Applets/FUSspecies.aspx" title="FUS Alignment and
Mutation"  description="" />
          <siteMapNode url="~/Chromosomes/chromo1.aspx" title="Chromosome
1"  description="" />
          <siteMapNode url="~/Chromosomes/chromo2.aspx" title="Chromosome
2"  description="" />
          <siteMapNode url="~/Chromosomes/chromo3.aspx" title="Chromosome
3"  description="" />
          <siteMapNode url="~/Chromosomes/chromo4.aspx" title="Chromosome
4"  description="" />
          <siteMapNode url="~/Chromosomes/chromo5.aspx" title="Chromosome
5"  description="" />
          <siteMapNode url="~/Chromosomes/chromo6.aspx" title="Chromosome
6"  description="" />
          <siteMapNode url="~/Chromosomes/chromo7.aspx" title="Chromosome
7"  description="" />
          <siteMapNode url="~/Chromosomes/chromo8.aspx" title="Chromosome
8"  description="" />
          <siteMapNode url="~/Chromosomes/chromo9.aspx" title="Chromosome
9"  description="" />
          <siteMapNode url="~/Chromosomes/chromo10.aspx" title="Chromosome
10"  description="" />
          <siteMapNode url="~/Chromosomes/chromo11.aspx" title="Chromosome
11"  description="" />
          <siteMapNode url="~/Chromosomes/chromo12.aspx" title="Chromosome
12"  description="" />
          <siteMapNode url="~/Chromosomes/chromo13.aspx" title="Chromosome
13"  description="" />
          <siteMapNode url="~/Chromosomes/chromo14.aspx" title="Chromosome
14"  description="" />
          <siteMapNode url="~/Chromosomes/chromo15.aspx" title="Chromosome
15"  description="" />
          <siteMapNode url="~/Chromosomes/chromo16.aspx" title="Chromosome
16"  description="" />
          <siteMapNode url="~/Chromosomes/chromo17.aspx" title="Chromosome
17"  description="" />
          <siteMapNode url="~/Chromosomes/chromo18.aspx" title="Chromosome
18"  description="" />
          <siteMapNode url="~/Chromosomes/chromo19.aspx" title="Chromosome
19"  description="" />
          <siteMapNode url="~/Chromosomes/chromo20.aspx" title="Chromosome
20"  description="" />
          <siteMapNode url="~/Chromosomes/chromo21.aspx" title="Chromosome
21"  description="" />
          <siteMapNode url="~/Chromosomes/chromo22.aspx" title="Chromosome
22"  description="" />
          <siteMapNode url="~/Chromosomes/chromo23.aspx" title="Chromosome
23"  description="" />
```

```
          <siteMapNode url="~/Chromosomes/chromo24.aspx" title="Chromosome
24"  description="" />
          <siteMapNode url="~/Chromosomes/chromoAll.aspx" title="All
Chromosomes"  description="" />
    </siteMapNode>
      <siteMapNode url="~/Statistics/analysis.aspx" title="Analysis"
description="">
        <siteMapNode url="~/Overview/interaction.aspx" title="Interactions"
description="" />
        <siteMapNode url="~/Statistics/statistics.aspx" title="Side-by-side
Comparison"  description="" />
        <siteMapNode url="~/Statistics/report.aspx" title="Summary Report"
description="" />
        <siteMapNode url="~/Statistics/credibility.aspx" title="Gene
Credibility Score"  description="" />
        <siteMapNode url="~/Statistics/pathogenicity.aspx"
title="Pathogenicity"  description="" />
      </siteMapNode>
      <siteMapNode url="~/index7.aspx" title="Summary"  description="" />
      <siteMapNode url="~/Overview/search.aspx" title="Search"
description="" />
      <siteMapNode url="~/summary/summary.aspx" title="Database Search"
description="">
        <siteMapNode url="~/subjects/index.aspx" title="Patient data"
description="" />
        <siteMapNode url="~/misc/Top10.aspx" title="Last 10 submissions"
description="" />
        <siteMapNode url="~/misc/diseaseDetails.aspx" title="Disease"
description="" />
        <siteMapNode url="~/misc/literature.aspx" title="Literature"
description="" />
        <siteMapNode url="~/misc/usefulLinks.aspx" title="Useful Links"
description="" />
      </siteMapNode>
      <siteMapNode url="~/Search/searchWeb.aspx" title="ALS News"
description="" >
          <siteMapNode url="~/BlogList.aspx" title="View comments"
description="" />
          <siteMapNode url="~/BlogEntry.aspx" title="Make a comment"
description="" />
        </siteMapNode>


      <siteMapNode url="~/charts/index.aspx" title="Visitors Statistics"
description="" />
      <siteMapNode url="~/GWA2/index.aspx" title="GWAS"  description="">
        <siteMapNode url="~/GWA2/whole.aspx" title="Whole Genome"
description="" />
        <siteMapNode url="~/GWA2/chisquare.aspx" title="On-the-fly
analysis"  description="" />
        <siteMapNode url="~/GWA2/gwa.aspx" title="meta-analysis"
description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo1.aspx" title="Chromosome
1"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo2.aspx" title="Chromosome
2"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo3.aspx" title="Chromosome
3"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo4.aspx" title="Chromosome
4"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo5.aspx" title="Chromosome
5"  description="" />
```

```
        <siteMapNode url="~/DataSets/Tracks/chromo6.aspx" title="Chromosome
6"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo7.aspx" title="Chromosome
7"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo8.aspx" title="Chromosome
8"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo9.aspx" title="Chromosome
9"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo10.aspx"
title="Chromosome 10"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo11.aspx"
title="Chromosome 11"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo12.aspx"
title="Chromosome 12"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo13.aspx"
title="Chromosome 13"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo14.aspx"
title="Chromosome 14"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo15.aspx"
title="Chromosome 15"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo16.aspx"
title="Chromosome 16"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo17.aspx"
title="Chromosome 17"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo18.aspx"
title="Chromosome 18"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo19.aspx"
title="Chromosome 19"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo20.aspx"
title="Chromosome 20"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo21.aspx"
title="Chromosome 21"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromo22.aspx"
title="Chromosome 22"  description="" />
        <siteMapNode url="~/DataSets/Tracks/chromoX.aspx" title="Chromosome
X"  description="" />
      </siteMapNode>
        <siteMapNode url="~/Overview/gene.aspx" title="Gene Overview"
description="" >
          <siteMapNode url="~/reports/allPatients.aspx" title="Patients in
database"  description="" />
          </siteMapNode>
      <siteMapNode url="~/Overview/protein.aspx" title="Protein Overview"
description="" />
        <siteMapNode url="~/Overview/study.aspx" title="Gene Study"
description="" />
        <siteMapNode url="~/misc/contributors.aspx" title="Contributors"
description="" >
        <siteMapNode url="~/registration/contributors_detail.aspx"
title="Contributors' details"  description="" />
        </siteMapNode>
      <siteMapNode url="~/misc/labs.aspx" title="Laboratories"
description="" />
        <siteMapNode url="~/contact/contact.aspx" title="Contact Us"
description="" />
        <siteMapNode url="~/misc/dataDownload.aspx" title="Mutation Data"
description="" />
        <siteMapNode url="~/contact/guestbook.aspx" title="Feedback"
description="" />
        <siteMapNode url="~/misc/FAQs.aspx" title="FAQs"  description="" />
      <siteMapNode url="~/misc/resources.aspx" title="Resources and Links"
description="" />
```

```
      <siteMapNode url="~/database/gene/credibilitySurveymonkey.aspx"
title="Credibility Survey by ALS Experts"  description="" />
        <siteMapNode url="~/mutations/mutationsFound.aspx"
title="Mutations"  description="" />
      <siteMapNode url="~/mutations/mutationsFoundCodon.aspx"
title="Mutations"  description="" />
      <siteMapNode url="~/mutations/mutationsFoundGeneOnly.aspx"
title="Mutations"  description="" />
      <siteMapNode url="~/mutations/mutationsFoundMnemonicOnly.aspx"
title="Mutations"  description="" />
      <siteMapNode url="~/mutations/mutationsFoundNoMnemonic.aspx"
title="Mutations"  description="" />
      <siteMapNode url="~/mutations/mutationsFoundNoType.aspx"
title="Mutations"  description="" />
      <siteMapNode url="~/mutations/mutationsFoundTypeOnly.aspx"
title="Mutations"  description="" />
      <siteMapNode url="~/results/showMutationDetails.aspx" title="Mutation
Details"  description="" />
      <siteMapNode url="~/results/showPatients.aspx" title="Patient
Details"  description="" />
      <siteMapNode url="~/login/loginAuthenticate.aspx" title="Login"
description="" />
      <siteMapNode url="~/database/patient/submitPatient000.aspx"
title="Submit patient data (1 of 5)"  description="" />
      <siteMapNode url="~/database/patient/submitPatient001.aspx"
title="Submit patient data (2 of 5)"  description="" />
      <siteMapNode url="~/database/patient/submitPatient002.aspx"
title="Submit patient data (3 of 5)"  description="" />
      <siteMapNode url="~/database/patient/submitPatient003.aspx"
title="Submit patient data (4 of 5)"  description="" />
      <siteMapNode url="~/database/patient/submitPatient004.aspx"
title="Submit patient data (5 of 5)"  description="" />
      <siteMapNode url="~/database/gene/credibilitySurvey.aspx"
title="Credibility Survey by ALS Experts"  description="" />
        <siteMapNode url="~/misc/Thankyou.aspx" title="Thank you"
description="" />
        <siteMapNode url="~/login/securePage.aspx" title="Login"
description="" >
          <siteMapNode
url="~/database/mutation/deletionEnterMutationDetails.aspx" title="Enter
details"  description="" />
          <siteMapNode
url="~/database/mutation/enterIntronMutationDocumentation.aspx"
title="Enter Documentation"  description="" />
          <siteMapNode
url="~/database/mutation/enterMutationDocumentation.aspx" title="Enter
Documentation"  description="" />
          <siteMapNode
url="~/database/mutation/enterMutationTypeLocation.aspx" title="Enter
Mutation Location"  description="" />
          <siteMapNode url="~/database/mutation/exonChooseCodon.aspx"
title="Choose Codon"  description="" />
          <siteMapNode
url="~/database/mutation/exonChooseNucleotidePosition.aspx" title="Choose
Nucleotide Position"  description="" />
          <siteMapNode
url="~/database/mutation/exonEnterMutationDetails.aspx" title="Enter
Details"  description="" />
          <siteMapNode url="~/database/mutation/exonMutationDetails.aspx"
title="Mutation Details"  description="" />
```

```
          <siteMapNode
url="~/database/mutation/intronEnterMutationDetails.aspx" title="Enter
Details"  description="" />
          <siteMapNode url="~/database/mutation/readyToSubmitMutation.aspx"
title="Submit"  description="" />
          <siteMapNode
url="~/database/mutation/verifyIntronMutationDetails.aspx" title="Verify
Details"  description="" />
          <siteMapNode url="~/database/mutation/verifyMutationDetails.aspx"
title="Verify Details"  description="" />
        </siteMapNode>
      <siteMapNode url="~/maps/mutationmap.aspx" title="Google Earth
Mutation Map"  description="" />
    </siteMapNode>
</siteMap>
```

- Blogs
- Social Media
  - Facebook
  - Twitter
- Data
  - Detailed Analysis
  - Interactions
  - Side-by-side comparison
  - Predict Pathogenicity
  - Gene Credibility
- Contributors
  - Detailed Analysis
  - Interactions
  - Side-by-side comparison
  - Predict Pathogenicity
  - Gene Credibility
- Feedback
  - Detailed Analysis
  - Interactions
  - Side-by-side comparison
  - Predict Pathogenicity
  - Gene Credibility
- Top section
  - Login
  - Feedback
  - Help
  - Register
  - Cookies
  - Print
  - Cite us
  - Sitemap
- Left section

- Left section
  - Chromosomes
    - Chromosome 1
    - Chromosome 2
    - Chromosome 3
    - Chromosome 4
    - Chromosome 5
    - Chromosome 6
    - Chromosome 7
    - Chromosome 8
    - Chromosome 9
    - Chromosome 10
    - Chromosome 11
    - Chromosome 12
    - Chromosome 13
    - Chromosome 14
    - Chromosome 15
    - Chromosome 16
    - Chromosome 17
    - Chromosome 18
    - Chromosome 19
    - Chromosome 20
    - Chromosome 21
    - Chromosome 22
    - Chromosome ALL
  - Funders
  - Visit Statistics
  - Facebook

Appendix 14 – Tracking visitors script

```
protected void Page_Load(object sender, EventArgs e)
   {


//Method for Mobile device detection
        DetectUserAgent();
        this.countMe();
  DataSet tmpDs = new DataSet();
        tmpDs.ReadXml(Server.MapPath("~/count/counter.xml"));
        lblCounter.Text = tmpDs.Tables[0].Rows[0]["hits"].ToString();


//Track Visitors' hostname
        string hostName =
System.Net.Dns.GetHostEntry(Request.ServerVariables["SERVER_NAME"]).HostName;


      //Convert IP Address to IP Number
        string ipAddress = GetIP4Address();
        string[] arrDec;
        UInt64 ipnum = 0;
        if (ipAddress != "")
        {
            arrDec = ipAddress.Split('.');
            ipnum = (UInt64.Parse(arrDec[3])) + (UInt64.Parse(arrDec[2]) * 256) +
(UInt64.Parse(arrDec[1]) * 65536) + (UInt64.Parse(arrDec[0]) * 16777216);
        }
             Label4.Text = ipnum.ToString();
//Get country of visitor
        String country;
        DataView dv2 = (DataView)SqlDataSource2.Select(DataSourceSelectArguments.Empty);
        if (dv2.Table.Rows.Count > 0)
        {
            country = (string)dv2.Table.Rows[0][0];
        }
        else
```

```
        {
            country = "Unknown";
        }
        Label7.Text = country;

        //Get country code of visitor
        String countrycode;
        DataView dv3 = (DataView)SqlDataSource2.Select(DataSourceSelectArguments.Empty);
        if (dv3.Table.Rows.Count > 0)
        {
            countrycode = (string)dv2.Table.Rows[0][1];
        }
        else
        {
            countrycode = "NIL";
        }
        Label8.Text = countrycode;

         //Write to visitors.log
        StreamWriter wrtr = new StreamWriter(Server.MapPath("~/visitors.log"), true);
        wrtr.WriteLine(DateTime.Now.ToString() + " | " + ipAddress + " | " + hostName + " | " + country +
" | " + Request.Url.ToString());
        wrtr.Close();

      //Write same data saved in visitor's log into database
    String DBconnection =
ConfigurationManager.ConnectionStrings["alsodConnectionString1"].ToString();
    SqlConnection con = new SqlConnection(DBconnection);
    SqlCommand cmd = new SqlCommand("INSERT INTO [visitors] (DateTime, IPAddress, Host,
Country, CountryCode, Page)VALUES(@DateTime, @IPAddress, @Host, @Country,
@CountryCode, @Page)", con);
    cmd.Parameters.Add(new SqlParameter("@DateTime", DateTime.Now.ToString("F")));
    cmd.Parameters.Add(new SqlParameter("@IPAddress", ipAddress));
    cmd.Parameters.Add(new SqlParameter("@Host", hostName));
    cmd.Parameters.Add(new SqlParameter("@Country", country));
```

```
    cmd.Parameters.Add(new SqlParameter("@CountryCode", countrycode));
    cmd.Parameters.Add(new SqlParameter("@Page", Request.Url.ToString()));
    cmd.Connection.Open();
    cmd.ExecuteNonQuery();
    cmd.Connection.Close();

        //Using Database to store counter
        DataView dv1 = (DataView)SqlDataSource1.Select(DataSourceSelectArguments.Empty);
        int count = (int)dv1[0][0];
        count = count + 1;
        SqlDataSource1.UpdateCommand = "UPDATE [COUNTER] SET [COUNT] = " + count;
        SqlDataSource1.Update();
        DataView users = ((DataView)SqlDataSource3.Select(DataSourceSelectArguments.Empty));
        Label9.Text = users.Table.Rows[0][0].ToString();

   }

    private string IpAddress()
    {
        string strIpAddress;
        strIpAddress = Request.ServerVariables["HTTP_X_FORWARDED_FOR"];
        if (strIpAddress == null)
            strIpAddress = Request.ServerVariables["REMOTE_ADDR"];
        return strIpAddress;
    }

    public static string GetIP4Address()
    {
    string IP4Address = String.Empty;
foreach (IPAddress IPA in Dns.GetHostAddresses(HttpContext.Current.Request.UserHostAddress))
    {
      if (IPA.AddressFamily.ToString() == "InterNetwork")
       {
```

```
    IP4Address = IPA.ToString();
    break;
  }
}

if (IP4Address != String.Empty)
{
  return IP4Address;
}

foreach (IPAddress IPA in Dns.GetHostAddresses(Dns.GetHostName()))
{
  if (IPA.AddressFamily.ToString() == "InterNetwork")
  {
    IP4Address = IPA.ToString();
    break;
  }
}

return IP4Address;
        }

private void countMe()
{
    DataSet tmpDs = new DataSet();
    tmpDs.ReadXml(Server.MapPath("~/count/counter.xml"));
    int hits = Int32.Parse(tmpDs.Tables[0].Rows[0]["hits"].ToString());
    hits += 1;
    tmpDs.Tables[0].Rows[0]["hits"] = hits.ToString();
    tmpDs.WriteXml(Server.MapPath("~/count/counter.xml"));
}
```

Appendix 15 – Inserting association text files into SQL Server

```
BULK INSERT dbo.[GWA_NOKEY]
FROM '\\vm3\Alsod\DataSets\uk\uk_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO



BULK INSERT dbo.GWA_NOKEY
FROM '\\vm3\Alsod\DataSets\bos\bos_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO


BULK INSERT dbo.GWA_NOKEY
FROM '\\vm3\Alsod\DataSets\hol\hol_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO


BULK INSERT dbo.GWA_NOKEY
FROM '\\vm3\Alsod\DataSets\fra\fra_result_assoc.txt'
```

```
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO


BULK INSERT dbo.GWA_NOKEY
FROM '\\vm3\Alsod\DataSets\usa\usa_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO
```

Appendix 16 – ALTER TABLE for inserting ID and population column

```
---Insert ID and Population column
ALTER TABLE dbo.GWA_NOKEY
ADD ID INT IDENTITY(1,1);
GO

ALTER TABLE dbo.GWA_NOKEY
ADD POP VARCHAR(50);
GO

UPDATE dbo.GWA_NOKEY SET POP = 'UK'
WHERE ID BETWEEN 1 AND 275619
GO

UPDATE dbo.GWA_NOKEY SET POP = 'BOSTON'
WHERE ID BETWEEN 275620 AND 558067
GO

UPDATE dbo.GWA_NOKEY SET POP = 'HOLLAND'
WHERE ID BETWEEN 558068 AND 846203
GO

UPDATE dbo.GWA_NOKEY SET POP = 'FRANCE'
WHERE ID BETWEEN 846204 AND 1132710
GO

UPDATE dbo.GWA_NOKEY SET POP = 'USA'
WHERE ID BETWEEN 1132711 AND 1420350
GO
```

Appendix 17 – Codes for creating scientific hyperlinks section

Entrez Gene [526] – An NCBI information

http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gene&Cmd=ShowDetailView&TermToSearch=6647&ordinalpos=3&itool=EntrezSystem2.PEntrez.Gene.Gene_ResultsPanel.Gene_RVDocSum

The codes on ASP.NET to generate this link are as shown below:

```
Dim ALSgene As String = http://www.alsgene.org/geneoverview.asp?geneid=

Dim dv20 As DataView

Dim ALSgene_id As String

dv20 = CType(SqlDataSource20.Select(DataSourceSelectArguments.Empty), DataView)

ALSgene_id = CType(dv20.Table.Rows(0)(0), String)

Label32.Text = ALSgene + ALSgene_id

HyperLink32.NavigateUrl = Label32.Text


'Shows ALSgene links available

    If (ALSgene_id <> 0) Then

        HyperLink32.Visible = "true"

    Else

       HyperLink32.Visible = "false"

    End If
```

```
<asp:HyperLink ID="HyperLink32" runat="server" Target="_blank" ToolTip="The ALSGene database will provide a comprehensive, unbiased and regularly updated field synopsis of genetic association studies performed in ALS.">ALSGene</asp:HyperLink>  

<asp:SqlDataSource ID="SqlDataSource20" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

    SelectCommand="SELECT [ALSgene_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

  <SelectParameters>

      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

    </SelectParameters>

  </asp:SqlDataSource>
```

UCSC Browser [527] – Genome browser

http://genome.ucsc.edu/cgi-bin/hgTracks?Submit=Submit&db=hg19&position=chr21:33,031,935-33,041,241

The codes on ASP.NET to generate this link are as shown below:

```
Dim UCSCDetails As String = “http://genome.ucsc.edu/cgi-bin/hgTracks?Submit=Submit&db=hg19&position=”

Dim dv2 As DataView

Dim UCSC As String

dv2 = CType(SqlDataSource4.Select(DataSourceSelectArguments.Empty), DataView)

    UCSC = CType(dv2.Table.Rows(0)(0), String)

Label2.Text = UCSCDetails + UCSC

HyperLink2.NavigateUrl = Label2.Text


<asp:HyperLink ID="HyperLink2" runat="server" Target="_blank" ToolTip="UCSC Genome Browser in Human">UCSC Browser</asp:HyperLink>  

<asp:SqlDataSource ID="SqlDataSource2" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

    SelectCommand="SELECT [Lastname], [Year],[gene_id],[Title],[PMID],[Paperlink] FROM dbo.gene_study WHERE ([gene_id] = @gene_id) UNION SELECT DISTINCT first_author, year, gene_id, title, pubmed_id, link FROM dbo.mutation WHERE (gene_id = @gene_id) AND (first_author <> 'NULL') UNION SELECT DISTINCT First_Author, Year, gene_id, paperlink, pubmed_id, link FROM dbo.gene_SNP_updated WHERE (First_Author <> 'NULL') AND (paperlink <> 'NULL') AND (gene_id = @gene_id)">

    <SelectParameters>

      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

    </SelectParameters>

  </asp:SqlDataSource>
```

Protein Structure[528] – Protein Domain and structure information as displayed on the UCSC Website

http://genome.cse.ucsc.edu/cgi-bin/hgGene?hgg_gene=uc002ypa.1#domains

The codes on ASP.NET to generate this link are as shown below:

```
Dim StructureDiagram As String = "http://genome.cse.ucsc.edu/cgi-bin/hgGene?hgg_gene="

Dim dv3 As DataView

Dim Diagram As String

dv3 = CType(SqlDataSource5.Select(DataSourceSelectArguments.Empty), DataView)

    Diagram = CType(dv3.Table.Rows(0)(0), String)

Label3.Text = StructureDiagram + Diagram + exactlocation

HyperLink3.NavigateUrl = Label3.Text


<asp:HyperLink ID="HyperLink3" runat="server" Target="_blank" ToolTip="3D structure of gene">Protein Structure</asp:HyperLink>  
```

<asp:SqlDataSource ID="SqlDataSource5" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

SelectCommand="SELECT [structure_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

<SelectParameters>

<asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

</SelectParameters>

</asp:SqlDataSource>

OMIM [529, 530] – It shows extensive information from the Online Mendelian Inheritance in Man database updated –though not regularly- by the research community.

http://www.ncbi.nlm.nih.gov/omim/147450

The codes on ASP.NET to generate this link are as shown below:

Dim OmimUrl As String = “http://www.ncbi.nlm.nih.gov/omim/”

Dim omim As String

Dim dv4 As DataView

dv4 = CType(SqlDataSource6.Select(DataSourceSelectArguments.Empty), DataView)

omim = CType(dv4.Table.Rows(0)(0), String)

Label4.Text = OmimUrl + omim

HyperLink4.NavigateUrl = Label4.Text

<asp:HyperLink ID="HyperLink4" runat="server" Target="_blank" ToolTip="Shows the Online Mendelian Inheritance In Man">OMIM</asp:HyperLink>

<asp:SqlDataSource ID="SqlDataSource6" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

SelectCommand="SELECT [omim_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

<SelectParameters>

<asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

</SelectParameters>

</asp:SqlDataSource>

Genecards[531] – These also show users an overview of gene information curated for research purpose.

http://www.genecards.org/cgi-bin/carddisp.pl?id=11179&id_type=hgnc&search=11179

The codes on ASP.NET to generate this link are as shown below:

Dim geneCardUrl1 As String = "http://www.genecards.org/cgi-bin/carddisp.pl?id="

```
Dim geneCardUrl2 As String = "&id_type=hgnc&search="

Dim dv5 As DataView

Dim hgnc As String

dv5 = CType(SqlDataSource7.Select(DataSourceSelectArguments.Empty), DataView)

hgnc = CType(dv5.Table.Rows(0)(0), String)

Label5.Text = geneCardUrl1 + hgnc + geneCardUrl2 + hgnc

HyperLink5.NavigateUrl = Label5.Text
```

```
<asp:HyperLink ID="HyperLink5" runat="server" Target="_blank" ToolTip="Shows the summary on gene">Genecards</asp:HyperLink>  
```

```
<asp:SqlDataSource ID="SqlDataSource7" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

    SelectCommand="SELECT [hgnc_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

  <SelectParameters>

      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

    </SelectParameters>

  </asp:SqlDataSource>
```

Full Literature – This displays the result of a pubmed query e.g. for SOD1 gene, “SOD1[All Fields] AND amyotrophic lateral sclerosis/genetics[mh] AND motor neuron disease/genetics[mh] AND "humans"[MeSH Terms] AND english[la]”

The codes on ASP.NET to generate this link are as shown below:

```
Dim literatureSearchUrl1 As String = "http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&term="

Dim literatureSearchUrl2 As String = "+AND+amyotrophic+lateral+sclerosis+/+genetics*[mh]+Motor+Neuron+Disease+/+genetics*[mh]+AND+humans[mh]AND+english[la]&db=PubMed&orig_db=PubMed&filters=on"

Dim dv7 As DataView

    Dim keywords As String

dv7 = CType(SqlDataSource9.Select(DataSourceSelectArguments.Empty), DataView)

    keywords = CType(dv7.Table.Rows(0)(0), String)

Label6.Text = literatureSearchUrl1 + keywords + literatureSearchUrl2

HyperLink6.NavigateUrl = Label6.Text
```

```
<asp:HyperLink ID="HyperLink6" runat="server" Target="_blank" ToolTip="Shows all available literature in pubmed on this gene">Full Literature</asp:HyperLink>  
```

```
<asp:SqlDataSource ID="SqlDataSource9" runat="server" ConnectionString="<%$
ConnectionStrings:alsodConnectionString1 %>"
    SelectCommand="SELECT [keywords] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

ProtScale [532] – Links to an external website programmed to automatically use the uniprot identifier for the selected gene. This enables a computation and representation of the profile generated by an amino acid scale on the chosen protein.

http://www.expasy.org/cgi-bin/protscale.pl?P00441

The codes on ASP.NET to generate this link are as shown below:

```
Dim protScaleUrl As String = “http://www.expasy.org/cgi-bin/protscale.pl?”
Dim dv8 As DataView
Dim protScale As String
dv8 = CType(SqlDataSource10.Select(DataSourceSelectArguments.Empty), DataView)
protScale = CType(dv8.Table.Rows(0)(0), String)
Label7.Text = protScaleUrl + protScale
HyperLink7.NavigateUrl = Label7.Text
```

```
<asp:HyperLink ID="HyperLink7" runat="server" Target="_blank" ToolTip="Shows the amino acid
scale">ProtScale</asp:HyperLink> 
```

```
<asp:SqlDataSource ID="SqlDataSource10" runat="server" ConnectionString="<%$
ConnectionStrings:alsodConnectionString1 %>"
    SelectCommand="SELECT [swissport_id] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

KEGG [533] – shows the AA and NT sequence from the KEGG database to save valuable time spent by researchers in searching for sequences of specific genes online.

http://www.genome.jp/dbget-bin/www_bget?hsa:6647

The codes on ASP.NET to generate this link are as shown below:

```
Dim sequenceUrl As String = http://www.genome.jp/dbget-bin/www_bget?hsa:

Dim dv9 As DataView

Dim sequence As String

dv9 = CType(SqlDataSource11.Select(DataSourceSelectArguments.Empty), DataView)

sequence = CType(dv9.Table.Rows(0)(0), String)

Label8.Text = sequenceUrl + sequence

HyperLink8.NavigateUrl = Label8.Text
```

```
<asp:HyperLink ID="HyperLink8" runat="server" Target="_blank" ToolTip="Shows the gene sequence in KEGG">Sequence</asp:HyperLink> 
```

```
<asp:SqlDataSource ID="SqlDataSource11" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

    SelectCommand="SELECT [ncbi_locuslink_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

  <SelectParameters>

      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

    </SelectParameters>

  </asp:SqlDataSource>
```

Uniprot [534] - The UniProt consortium provides access to the multiple data sets from the Swiss Institute of Bioinformatics (SIB), the European Bioinformatics Institute (EBI) and the Protein Information Resource (PIR).

http://www.uniprot.org/uniprot/P00441

The codes on ASP.NET to generate this link are as shown below:

```
Dim uniprotUrl As String = "http://www.uniprot.org/uniprot/"

Dim dv8 As DataView

Dim protScale As String

dv8 = CType(SqlDataSource10.Select(DataSourceSelectArguments.Empty), DataView)

protScale = CType(dv8.Table.Rows(0)(0), String)

Label10.Text = uniprotUrl + protScale

HyperLink10.NavigateUrl = Label10.Text
```

```
<asp:HyperLink ID="HyperLink10" runat="server" ToolTip="Shows the summary of gene in Uniprot" Target="_blank">Uniprot</asp:HyperLink>
```

```
<asp:SqlDataSource ID="SqlDataSource10" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"
```

```
    SelectCommand="SELECT [swissport_id] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

iHop [535] – Information Hyperlinked over Proteins database is a vastly connected network making human literature research more intuitive and efficient using a text-mining method as described in their publication. This is linked from ALSoD to allow researchers an unbiased list of published materials relating to the selected gene.

http://www.ihop-net.org/UniPub/iHOP/bng/92317.html?ORGANISM_ID=1

The codes on ASP.NET to generate this link are as shown below:

```
Dim ihopUrl1 As String = "http://www.ihop-net.org/UniPub/iHOP/bng/"
Dim ihopUrl2 As String = ".html?ORGANISM_ID=1"
Dim dv11 As DataView
Dim ihop As String
dv11 = CType(SqlDataSource13.Select(DataSourceSelectArguments.Empty), DataView)
    ihop = CType(dv11.Table.Rows(0)(0), String)
Label11.Text = ihopUrl1 + ihop + ihopUrl2
HyperLink11.NavigateUrl = Label11.Text
```

```
<asp:HyperLink ID="HyperLink11" runat="server" Target="_blank" ToolTip="Shows the latest information">iHop</asp:HyperLink>  
```

```
<asp:SqlDataSource ID="SqlDataSource13" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"
    SelectCommand="SELECT [ihop] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

Pathway in KEGG [533] – A KEGG pathway map of human Amyotrophic Lateral Sclerosis under the Neurodegenerative diseases of the Human disease section is hyperlinked from our website to display a pathway overview map.

http://www.genome.jp/dbget-bin/show_pathway?hsa05014

The codes on ASP.NET to generate this link are as shown below:

Dim pathway As String = "http://www.genome.jp/dbget-bin/show_pathway?hsa05014"

Label12.Text = pathway

HyperLink12.NavigateUrl = Label12.Text

<asp:HyperLink ID="HyperLink12" runat="server" Target="_blank" ToolTip="Shows the pathway of ALS">Pathway</asp:HyperLink>

GeneTest *[536]* - The GeneTests database and Web site which is hosted at NCBI displays gene reviews and information on laboratories and clinics related to the gene. http://www.ncbi.nlm.nih.gov/sites/GeneTests/lab/research_disease_id/229174?db=genetests&report=Full

The codes on ASP.NET to generate this link are as shown below:

Dim genetest As String = "http://www.ncbi.nlm.nih.gov/sites/GeneTests/lab/research_disease_id/229174?db=genetests&report=Full"

Label19.Text = genetest

HyperLink19.NavigateUrl = Label19.Text

<asp:HyperLink ID="HyperLink19" runat="server" Target="_blank" ToolTip="Shows related genes on ALS in the geneTest webpage">GeneTest</asp:HyperLink>

ALS review – A link to the bookshelf page of NCBI displaying regularly reviewed information on ALS overview to enable researchers view up-to-date research in the field of ALS.

AmiGO *[537]* – A link to this free open source web application developed and kept up-to-date by the GO Consortium lets users query, browse and view ontology and annotation data on definite genes.

http://amigo.geneontology.org/cgi-bin/amigo/gp-assoc.cgi?gp=UniProtKB:P00441

The codes on ASP.NET to generate this link are as shown below:

Dim Amigo As String = "http://amigo.geneontology.org/cgi-bin/amigo/gp-assoc.cgi?gp=UniProtKB:"

dv8 = CType(SqlDataSource10.Select(DataSourceSelectArguments.Empty), DataView)

protScale = CType(dv8.Table.Rows(0)(0), String)

Label24.Text = Amigo + protScale

HyperLink24.NavigateUrl = Label24.Text

```
<asp:HyperLink ID="HyperLink24" runat="server" Target="_blank" ToolTip="Gene Ontology">AmiGO</asp:HyperLink>  

<asp:SqlDataSource ID="SqlDataSource10" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"
    SelectCommand="SELECT [swissport_id] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

Ensembl [538] – This displays a broad genome information system featuring an amalgamated set of genome annotation, databases, and other information.

http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000142168

The codes on ASP.NET to generate this link are as shown below:

```
Dim Ensemblurl As String = "http://www.ensembl.org/Homo_sapiens/Gene/Summary?g="
Dim ensembl As String
Dim dv17 As DataView
dv17 = CType(SqlDataSource17.Select(DataSourceSelectArguments.Empty), DataView)
    ensembl = CType(dv17.Table.Rows(0)(0), String)
Label27.Text = Ensemblurl + ensembl
HyperLink27.NavigateUrl = Label27.Text

<asp:HyperLink ID="HyperLink27" runat="server" Target="_blank" ToolTip="Ensembl">Ensembl</asp:HyperLink>  

<asp:SqlDataSource ID="SqlDataSource17" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"
    SelectCommand="SELECT [ensembl_id] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

NCBI [480] – This comprehensive database and website is linked from ALSoD showing a statistical overview of the dataset across its network of databases e.g. for

ANG gene, the search query automatically coded to our database is “ANG HOMO ALS” displaying a total of 31 pubmed citations and abstracts.

http://www.ncbi.nlm.nih.gov/sites/gquery?term=SOD1%20HOMO%20ALS

The codes on ASP.NET to generate this link are as shown below:

```
Dim ncbi1 As String = "http://www.ncbi.nlm.nih.gov/sites/gquery?term="
Dim ncbi2 As String = "%20HOMO%20ALS"
Dim dv7 As DataView
Dim keywords As String
dv7 = CType(SqlDataSource9.Select(DataSourceSelectArguments.Empty), DataView)
keywords = CType(dv7.Table.Rows(0)(0), String)
Label29.Text = ncbi1 + keywords + ncbi2
HyperLink29.NavigateUrl = Label29.Text
```

```
<asp:HyperLink ID="HyperLink29" runat="server" Target="_blank" ToolTip="NCBI database summary of this gene">NCBI</asp:HyperLink>  
```

```
<asp:SqlDataSource ID="SqlDataSource9" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"
    SelectCommand="SELECT [keywords] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

Life Science DB (Japan) *[485]* – A recently developed ALS mutation database constructed as part of the Life Science Integrated Database Project conducted by the Japan Ministry of Education, Culture, Sports, Science, and Technology was published. It contains their original experimental results and published data extracted from scientific journals. The database is expected to play a complementary role to the ALSoD database especially in collecting variations in the Asian region. A link to this website allows a wider perspective of mutations available in ALS without geographical bias.

http://reseq.lifesciencedb.jp/resequence/GeneDetail.do?targetId=1&geneId=EG6647

The codes on ASP.NET to generate this link are as shown below:

```
Dim lifesciencedb As String = "http://reseq.lifesciencedb.jp/resequence/GeneDetail.do?targetId=1&geneId=EG"

dv9 = CType(SqlDataSource11.Select(DataSourceSelectArguments.Empty), DataView)

sequence = CType(dv9.Table.Rows(0)(0), String)

Label31.Text = lifesciencedb + sequence


'Shows Lifesciencedb (Japan) links available

If ((geneid = "SOD1") Or (geneid = "TARDBP(TDP43)") Or (geneid = "FUS") Or (geneid = "VAPB") Or (geneid = "DCTN1") Or (geneid = "ALS2") Or (geneid = "CNTF") Or (geneid = "DPP6") Or (geneid = "SETX") Or (geneid = "CHMP2B") Or (geneid = "ANG") Or (geneid = "HFE") Or (geneid = "APEX1") Or (geneid = "APOE") Or (geneid = "ITPR2") Or (geneid = "LIF") Or (geneid = "MAOB") Or (geneid = "NEFH") Or (geneid = "PON1") Or (geneid = "PON2") Or (geneid = "PON3") Or (geneid = "FGGY") Or (geneid = "PRPH") Or (geneid = "SLC1A2") Or (geneid = "SMN1") Or (geneid = "SMN2") Or (geneid = "SPAST") Or (geneid = "VEGFA")) Then


            HyperLink31.Visible = "true"

      Else

         HyperLink31.Visible = "false"

      End If


<asp:HyperLink ID="HyperLink31" runat="server" Target="_blank" ToolTip="An ALS mutation database supported by the Ministry of Education, Culture, Sports, Science and Technology in Japan">Life Science DB(Japan)</asp:HyperLink>  

<asp:SqlDataSource ID="SqlDataSource11" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

      SelectCommand="SELECT [ncbi_locuslink_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

   <SelectParameters>

         <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

      </SelectParameters>

   </asp:SqlDataSource>
```

ALSGene [539] – “ALSGene freely available on http://www.alsgene.org aims to serve as an exhaustive, unbiased, and regularly updated resource of genetic association studies in ALS. One of its key features will be up-to-date meta-analyses of all eligible genetic polymorphisms that have been investigated for association with ALS risk”.

http://www.alsgene.org/geneoverview.asp?geneid=1

The codes on ASP.NET to generate this link are as shown below:

```
Dim ALSgene As String = http://www.alsgene.org/geneoverview.asp?geneid=

Dim dv20 As DataView
```

```
Dim ALSgene_id As String

dv20 = CType(SqlDataSource20.Select(DataSourceSelectArguments.Empty), DataView)

ALSgene_id = CType(dv20.Table.Rows(0)(0), String)

Label32.Text = ALSgene + ALSgene_id


'Shows ALSgene links available

    If (ALSgene_id <> 0) Then

        HyperLink32.Visible = "true"

    Else

       HyperLink32.Visible = "false"

    End If
```

```
<asp:HyperLink ID="HyperLink32" runat="server" Target="_blank" ToolTip="The ALSGene database will provide a comprehensive, unbiased and regularly updated field synopsis of genetic association studies performed in ALS.">ALSGene</asp:HyperLink>  

<asp:SqlDataSource ID="SqlDataSource20" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

    SelectCommand="SELECT [ALSgene_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

  <SelectParameters>

      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

    </SelectParameters>

  </asp:SqlDataSource>
```

Gen2Phen

http://www.gen2phen.org/gene/SOD1

The codes on ASP.NET to generate this link are as shown below:

```
Dim Gen2Phen As String ="http://www.gen2phen.org/gene/"

Dim dv6 As DataView

Dim geneid As Stringdv6 = CType(SqlDataSource8.Select(DataSourceSelectArguments.Empty), DataView)

geneid = CType(dv6.Table.Rows(0)(0), String)

Label39.Text = Gen2Phen + geneid

HyperLink39.NavigateUrl = Label39.Text
```

```
<asp:HyperLink ID="HyperLink39" runat="server" Target="_blank" ToolTip="GEN2PHEN project is a holistic approach to genotype-to-phenotype data.">Gen2Phen</asp:HyperLink>  
```

<asp:SqlDataSource ID="SqlDataSource8" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

SelectCommand="SELECT [gene_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

<SelectParameters>

<asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

</SelectParameters>

</asp:SqlDataSource>

Orphanet

http://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=en&Gene=11179

The codes on ASP.NET to generate this link are as shown below:

Dim Orphanet As String ="http://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=en&Gene="

Dim dv5 As DataView

Dim hgnc As String

dv5 = CType(SqlDataSource7.Select(DataSourceSelectArguments.Empty), DataView)

hgnc = CType(dv5.Table.Rows(0)(0), String)

Label41.Text = Orphanet + hgnc

HyperLink41.NavigateUrl = Label41.Text

<asp:HyperLink ID="HyperLink32" runat="server" Target="_blank" ToolTip="The ALSGene database will provide a comprehensive, unbiased and regularly updated field synopsis of genetic association studies performed in ALS.">ALSGene</asp:HyperLink>

<asp:SqlDataSource ID="SqlDataSource7" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

SelectCommand="SELECT [hgnc_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

<SelectParameters>

<asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

</SelectParameters>

</asp:SqlDataSource>

GeneNetwork

http://129.125.165.109:8080/Expression/?gene=SOD1

The codes on ASP.NET to generate this link are as shown below:

```
Dim ALSgene As String = http://www.alsgene.org/geneoverview.asp?geneid=
```

```
Dim dv20 As DataView

Dim ALSgene_id As String

dv20 = CType(SqlDataSource20.Select(DataSourceSelectArguments.Empty), DataView)

ALSgene_id = CType(dv20.Table.Rows(0)(0), String)

Label32.Text = ALSgene + ALSgene_id


'Shows ALSgene links available

    If (ALSgene_id <> 0) Then

        HyperLink32.Visible = "true"

    Else

       HyperLink32.Visible = "false"

    End If
```

```
<asp:HyperLink ID="HyperLink32" runat="server" Target="_blank" ToolTip="The ALSGene database will provide a comprehensive, unbiased and regularly updated field synopsis of genetic association studies performed in ALS.">ALSGene</asp:HyperLink>  

<asp:SqlDataSource ID="SqlDataSource20" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"

    SelectCommand="SELECT [ALSgene_id] FROM [gene] WHERE ([gene_id] = @gene_id)">

  <SelectParameters>

      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />

    </SelectParameters>

  </asp:SqlDataSource>
```

MGI

http://www.informatics.jax.org/marker/MGI:98351

The codes on ASP.NET to generate this link are as shown below:

```
Dim ALSgene As String = http://www.alsgene.org/geneoverview.asp?geneid=

Dim dv20 As DataView

Dim ALSgene_id As String

dv20 = CType(SqlDataSource20.Select(DataSourceSelectArguments.Empty), DataView)

ALSgene_id = CType(dv20.Table.Rows(0)(0), String)

Label32.Text = ALSgene + ALSgene_id


'Shows ALSgene links available
```

```
    If (ALSgene_id <> 0) Then
        HyperLink32.Visible = "true"
    Else
       HyperLink32.Visible = "false"
    End If
```

```
<asp:HyperLink ID="HyperLink32" runat="server" Target="_blank" ToolTip="The ALSGene database will provide a comprehensive, unbiased and regularly updated field synopsis of genetic association studies performed in ALS.">ALSGene</asp:HyperLink>  

<asp:SqlDataSource ID="SqlDataSource20" runat="server" ConnectionString="<%$ ConnectionStrings:alsodConnectionString1 %>"
    SelectCommand="SELECT [ALSgene_id] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

LOVD

http://www.lovd.nl/SOD1

The codes on ASP.NET to generate this link are as shown below:

```
Dim ALSgene As String = http://www.alsgene.org/geneoverview.asp?geneid=

Dim dv20 As DataView

Dim ALSgene_id As String

dv20 = CType(SqlDataSource20.Select(DataSourceSelectArguments.Empty), DataView)

ALSgene_id = CType(dv20.Table.Rows(0)(0), String)

Label32.Text = ALSgene + ALSgene_id


'Shows ALSgene links available

    If (ALSgene_id <> 0) Then
        HyperLink32.Visible = "true"
    Else
       HyperLink32.Visible = "false"
    End If
```

```
<asp:HyperLink ID="HyperLink32" runat="server" Target="_blank" ToolTip="The ALSGene database will provide a comprehensive, unbiased and regularly updated field synopsis of genetic association studies performed in ALS.">ALSGene</asp:HyperLink>  
```

```
<asp:SqlDataSource ID="SqlDataSource20" runat="server" ConnectionString="<%$
ConnectionStrings:alsodConnectionString1 %>"
    SelectCommand="SELECT [ALSgene_id] FROM [gene] WHERE ([gene_id] = @gene_id)">
  <SelectParameters>
      <asp:QueryStringParameter Name="gene_id" QueryStringField="gene_id" Type="String" />
    </SelectParameters>
  </asp:SqlDataSource>
```

Appendix 18 - Claustalw and Jalview codes

```
Dim dv6 As DataView
Dim geneid As String
dv6     =     CType(SqlDataSource8.Select(DataSourceSelectArguments.Empty),
DataView)
geneid = CType(dv6.Table.Rows(0)(0), String)
Dim specie1 As String = "~/Applets/"
Dim specie2 As String = "species.aspx"
Label23.Text = specie1 + geneid + specie2
HyperLink23.NavigateUrl = Label23.Text
```

Where

```
<asp:SqlDataSource ID="SqlDataSource8" runat="server" ConnectionString="<%$
ConnectionStrings:alsodConnectionString1 %>"
SelectCommand="SELECT [gene_id] FROM [gene] WHERE ([gene_id] = @gene_id)">
    <SelectParameters>
<asp:QueryStringParameter     Name="gene_id"     QueryStringField="gene_id"
Type="String" />
        </SelectParameters>
    </asp:SqlDataSource>
```

The hyperlink redirects to an applet code on different species:

```
<applet code="jalview.bin.JalviewLite"
                              archive="jalviewApplet.jar"     style="width:
735px; height: 242px; margin-top: 0px;">
                        <param name="embedded" value="true">
                        <param                              name="sequence1"
value="P00441|SODC_HUMAN(Homo_sapiens)/0-153                    ---------
-----MATKAVCVLK-GDGP-VQGIINFEQK--ESNGPVKVWGSIKGLTE-GLHGF-HVHEFGDNTAG---
CTSAGPHFNPLSRKHGGPKDEERHVGDLGNVTADKDGVADVSIEDSVISLSGDH---
CIIGRTLVVHEKADDLGKGGNE---ESTKTGNAGSRLACGVIGIAQ-----------------------------
-">
                        <param                              name="sequence2"
value="Q95ME5|Q95ME5_PIG(Sus_scrofa)                           ---------
--------KAVCVLK-GDGP-VQGTIYFELK--GEK-TVLVTGTIKGLAE-GDHGF-HVHQFGDNTQG---
CTSAGPHFNPESKKHGGPKDQERHVGDLGNVTAGKDGVATVYIEDSVIALSGDH---
SIIGRTMVVHEKPDDLGRGGNE---ESTKTGNAGSRLACGVIG--------------------------------
-">
                        <param                              name="sequence3"
value="Q6LDS4|Q6LDS4_RAT(Rattus_norvegicus)                     ---------
-------MKAVCVLK-GDGP-VQGVIHFEQK--ASGEPVVVSGQITGLTE-GEHGF-HVHQYGDNTQG---
CTTAGPHFNPHSKKHGGPADEERHVGDLGNVAAGKDGVANVSIEDRVISLSGEH---
SIIGRTMVVHEKQDDLGKGGNE---ESTKTGNAGSRLACGVIGIAQ-----------------------------
-">
                        <param                              name="sequence4"
value="Q802E0|Q802E0_MELUD(Budgerigar)                          ---------
----MATLKAVCVMK-GEGP-VQGVIHFQQQ--GNG-PVKVTGKISGLAD-GDHGF-HVHEFGDNTNG---
CTSAGPHFNPEGKQHGGPSDAERHVGDLGNVTA-KGGVAEVAIEDSIISLSGPH---
SIVGRTMVVHEKCDDLGRGGDN---ESKLTGNAGPRLACGVIGIAKS----------------------------
-">
                        <param                              name="sequence5"
value="Q8QHI0|Q8QHI0_ONCMY(Rainbow_trout)                       ---------
-----MAMKAVCVLK-GTGE-VTGTVFFEQE--GADGPVKLIGEISGLAP-GEHGF-HVHAYGDNTNG---
CMSAGPHFNPHNQTHGGPTDAVRHVGDLGNVTAGADNVAKINIQDKMLTLTGPD---
SIIGRTMVIHEKADDLGKGGNE---ESLKTGNAGGRQACGVIGIAQ-----------------------------
-">
                        <param                              name="sequence6"
value="Q7YXM6|Q7YXM6_APILI(Common_honeybee)                     ---------
```

```
------MTKAVCVLQ--GE--VKGTIFFEQP--ESTNSVKVTGQVTGLKK-GLHGF-HVHEFGDNTNG---
CTSAGAHFNPLGKDHGGPDSDIRHVGDLGNIEADASGVANVNITDKTIQLQGPH---
SVIGRTLVVHADPDDLGKGGVE---LSKTTGNAGARLACGVIGITKV----------------------------
-">
                        <param                                  name="sequence7"
value="Q6W5R8|Q6W5R8_LASNI(Black_garden_ant)                          ---------
-----MTVKAVCVLQ--GEP-VKGTVHFEQA--DGSSAVKVTGEVSGLQK-GLHGF-HVHEFGDNTNG---
CTSAGAHFNPLGKEHGGPEHAVRHIGDLGNVEAGADGVAKINITDSQIQLSGPH---
SVIGRTVVVHADPDDLGQGGHE---LSKTTGNAGARLACGVIGITK----------------------------
-">
                        <param                                  name="sequence8"
value="Q9ZNQ4|Q9ZNQ4_CICAR(Chickpea)                                  ---------
------MVKAVAVLG--SSDTVSGTINFSQ---EGDGPTTVTGNLAGLKP-GLHGF-HIHALGDTTNG---
CISTGPHFNPNGKEHGSPEDPIRHAGDLGNINVGDDGTVSFSITDNQIPLTGPN---
SIIGRAVVVHADPDDLGKGGHE---LSKTTGNAGGRVACGIIGLQG----------------------------
-">
                        <param                                  name="sequence9"
value="Q70CE4|Q70CE4_FAGSY(Beechnut)                                   -
SVPTARGVLRSHRTMAKGVAVLS--SNEGVCGTIYFAQ---EGDGPTTVTGNISGLKP-GLHGF-
HVHALGDTTNG---CMSTGPHFNPAGKEHGAPEDANRHAGDLGNVNVGDDGTVSFTIIDKQIPLCGPN---
SIIGRAVVVHGDPDDLGKGGHE---LSKSTGNAGGRIACGIIGLQG----------------------------
-">
                        <param                                 name="sequence10"
value="Q9AR78|Q9AR78_9ROSI(Populus_tremuloides)                       ---------
------MVKAVAVLN--SSEGVKGTINFTQ---EGDGPTTVTGSLCGLKP-GLHGF-HVHALGDTTNG---
CMSTGPHFNPVGKEHGAPEDENRHAGDLGNVTVGDDGTATVSIIDNQIPLTGPN---
SIVGRAVVVHADPDDLGKGGHE---LSKSTGNAGGRVACGVIGLQG----------------------------
-">
                        <param                                 name="sequence11"
value="Q9FVF5|Q9FVF5_OLEEU(Common_olive)                              ---------
-------------------------------------------------------------FNG---
CMSTGPHFNPVGKEHGAPGDENRHAGDLGNITVGEDGTAAINIVDKQIPLTGPH---
SIIGRAVVVHSDPDDLGRGGHE---LSKRTGNAGGRTG-------------------------------------
-">
                        <param                                 name="sequence12"
value="Q9AR76|Q9AR76_9ROSI(Populus_tremuloides)                       ---------
---MATGSVKAVALIT-GDSNVRGSLHFIQ---EPNGATHVTGRITGLSP-GLHGF-HIHALGDTTNG---
CNSTGPHFNPLKKDHGAPSDNERHAGDLGNITAGSDGVAEVSIKDLQIPLSGMH---
SILGRAVVVHADPDDLGKGGHE---LSKTTGNAGARVGCGIIGLKSSV----------------------------
-">
                        <param                                 name="sequence13"
value="Q9Y0A5|Q9Y0A5_ACAVI(Filarial_nematode_worm)                    ---------
-------MSTNAIAVLR-GNTVSGVIRFKQD--KEGSPTIINGEIKGLTP-GLHGF-HIHQYGDTTNG---
CISAGPHFNPHNKTHGGPTDEIRHVGDLGNIVAGADGTAHIDIPNKQVQLLGPN---
SIIGRSIVVHADEDDLGKGVGDKKNESLKTGNAGARVACGIVAIGADS----------------------------
-">
                        <param                                 name="sequence14"
value="Q9C0S4|Q9C0S4_CRYNE(Cryptococcus_neoformans)                   ---------
-------MVKAVVVLK-GESYVHGTVCFTQE-SENAP-VCITGEIKDMDADAKRGM-HVHEFGDNTNG---
CTSAGPHYNPFKKHHGAPTDSERHVGDLGNIQTNSCGAAQLDFSDKIISLYGPH---
SIIGRSLVVHASTDDLGKGGNE---ESLKTGNAGARLACGVIGIST----------------------------
-">
                        <param                                 name="sequence15"
value="Q9C402|Q9C402_CRYNV(Filobasidiella_neoformans_var_grubii)          ---
-------------MVKAVVVLK-GESYAHGIVCFTQE-SENAP-VCITGEIKDMDADAKRGM-HVHEFGDNTNG-
--CTSAAPHYNPFKKHHGAPTDSERHVGDLGNIQTNSCGAAQLDFSDKIISLYGPH---
SIIGGSFVVHASTDDLGKGGNE---
ESLKTGNAGARLACGVIGISTCQCYHSKLIVFAAVFLPKRTVTTYSWLNK-">
                        <param                                 name="sequence16"
value="Q9C0T2|Q9C0T2_CRYNE(Cryptococcus_neoformans_var_neoformans)      ----
------------MVKAVAVLK-GDSHVYGTITFTQD-SEGAP-VCVSGEIKNLDADAKRGF-HVHEFGDNTNG--
-CTSAGPHYNPFHKNHGGPTAAERHVGDLGNVQTNGCGVAMVDISDKVISLFGPH---
```

```
SIIGRSMVVHAGTDDLGKGGNE---ESLKTGNAGARLACGVIGIAA-----------------------------
-">
                        <param                                name="sequence17"
value="Q6VTE9|Q6VTE9_CRYGA(Cryptococcus_bacillisporus)                  ---
----------------AVAVLK-GDSPVTGVITFTQE-KEGAP-VTVSGDIKNLDANAERGF-HVHEFGDNTNG-
--CTSAGPHFNPHGKNHGAPSDSERHVGDLGNVKTDSNGVASVNISDKSLSLFGPY---
SIIGRTIVVHAGTDDFGKGGNA---ESLKTGNAGARAACGV-----------------------------------
-">
                        <param                                name="sequence18"
value="Q96V80|Q96V80_9SACH(Debaryomyces_vanrijiae_var_vanrijiae)       ----
------------MVQPVTVFKRVYSKVIGVVNFEQS-SESDP-ISITWEISGNDANALIGF-HVHTFGDNTNG--
-CTSAGPHFNPFTKEHGAPEDDNRHVGDLGNVTTDTSGVAKGSKQDLFVKLIGQN---
SILGRTVVIHAGTDDLGKGGNA---ESKKTGNAGARLACGVIGLTN-----------------------------
-">
                        <param                                name="sequence19"
value="Q96WH8|Q96WH8_DEBHA(Yeast)                                ----------
-MVQAVAVLR-GDSKVIGVVNFEQS-SESDP-TFITWEISGNDANALRGF-HVHTFGDNTNG---
CTSAGPHFNPFTKEHGAPEDDNRHVGDLGNVTTDTSGVAKGSKQDLFVKLIGQN---
SILGRTVVIHAGTDDLGKGGNA---ESKKTGNAGARLACGVIGLTNKPNS-------------------------
-">
                        <param                                name="sequence20"
value="Q96V79|Q96V79_9TREE(Udeniomyces_puniceus)                         ---
-------------MVQAVVVLT-GDSMVTGVVNFGTNRSKSEPKLLFTWEISGNDANALKRFPTVHTFGDNTNG-
--CTSAGPHFNPFTKEHGAPEDDNRHVGDLGNVTTDTSGGAKGSKQDLFVKLIGQN---
SILGRTVVIHAGTDDLGKGGNA---ESKKTGNAGARLACGVIGLTN-----------------------------
-">
                        <param                                name="sequence21"
value="Q96WH7|Q96WH7_9SACH(Debaryomyces_vanrijiae_var_yarrowii)         ---
------------MVQAVAVLR--SDSKVS-RCRLTSNNRQSLTQLLLHGRFLGNDANALRGF-
HVHTFGDNTNGLYFCWTS-LHFNPFTKEHGAPEDDNRHVGDLGNVTTDTSGVAKGSKQDLFVKLIGQN---
SILGRTVVIHAGTDDLGKGGNA---ESKKTGNAGARLACGVIGLTNKPNS-------------------------
-">
                        <param                                name="sequence22"
value="Q8C355|Q8C355_MOUSE(Mus_musculus)                          ---------
-----MAMKAVCVLK--GDGPVQGTIHFEQK------------ARPGARG-AGRGD-AAHLCGSTPRP-------
RHGLSPLSAESPWPGAGAGRREARPGAPRGLPGGRGPPRAPERLGLPGRAGLASVILLGLLLFRCP---
CPTGSE---------PRGHRLRLL-------------------------------------">
                        <param name="defaultColour" value="Zappo">
                        <param name="showAnnotation" value="false">
                        <param name="windowHeight" value="515">
                        <param name="windowWidth" value="650">
                        <param name="showConservation" value="false">
                        <param name="showQuality" value="false">
                        <param name="showConsensus" value="false">

                        <param name="showFullId" value="false">
                        <param name="RGB"  value="F2F2FF">
                        <param name="linkLabel_1" value="SRS">
                        <param                                 name="linkUrl_1"
value="http://srs.ebi.ac.uk/srs7bin/cgi-bin/wgetz?-e+[uniprot-
all:$SEQUENCE_ID$]+-vn+2">
                        <param name="linkLabel_2" value="Uniprot">
                        <param                                 name="linkUrl_2"
value="http://us.expasy.org/cgi-bin/niceprot.pl?$SEQUENCE_ID$">

                    </applet>
```

The second section is an applet code on the variants of a selected gene:

```
<applet code="jalview.bin.JalviewLite"
```

```
                                archive="jalviewApplet.jar"    class="style1"
style="width: 734px; height: 168px">
                          <param name="embedded" value="true">
                          <param   name="sequence1"   value="SOD1_HUMAN/1-153
ATKAVCVLKGDGPVQGIINFEQKESNGPVKVWGSIKGLTEGLHGFHVHEFGDNTAGCTSAGPHFNPLSRKHGGPK
DEERHVGDLGNVTADKDGVADVSIEDSVISLSGDHCIIGRTLVVHEKADDLGKGGNEESTKTGNAGSRLACGVIG
IAQ" >
                          <param  name="sequence2"  value="ALSOD_LEVEL1  ---
SLFEQ-G-R-G-A--SCGL------A-------RR--D-R-CRAQK----R----I-----SAR----C---V--
-R---FRDA-TA--A-TNL-VGG--FLF-V--YMFAGTVL-----GH*R-----VN--REKA*--FG*DGT-S--
" >
                          <param name="sequence3" value="ALSOD_LEVEL2 ---T-G-
V-R---M-S----K----------------V--S-----FR----------S------------S---Y------
-VSIM-VV--C-VVM--KH----V-----TT----------V--------------NGE--STRRI--T--" >
                          <param name="sequence4" value="ALSOD_LEVEL3 ---V-S-
--V-------------------------------------------------------------------------
---K------D-------N-------------------------------------------------------" >
                          <param name="sequence5" value="ALSOD_LEVEL4 -------
-------------------------------------------------------------------------
---S------R-------Y-------------------------------------------------------" >
                          <param name="sequence6" value="ALSOD_LEVEL5 -------
-------------------------------------------------------------------------
----------S-------------------------------------------------------------">
                          <param name="sequence7" value="ALSOD_LEVEL6 -------
-------------------------------------------------------------------------
----------V-------------------------------------------------------------">
                          <param name="defaultColour" value="Zappo">
                          <param name="showAnnotation" value="false">
                          <param name="windowHeight" value="515">
                          <param name="windowWidth" value="650">
                          <param name="showConservation" value="false">
                          <param name="showQuality" value="false">
                          <param name="showConsensus" value="false">

                          <param name="showFullId" value="false">
                          <param name="RGB"  value="F2F2FF">
                          <param name="linkLabel_1" value="SRS">
                          <param                              name="linkUrl_1"
value="http://srs.ebi.ac.uk/srs7bin/cgi-bin/wgetz?-e+[uniprot-
all:$SEQUENCE_ID$]+-vn+2">
                          <param name="linkLabel_2" value="Uniprot">
                          <param                              name="linkUrl_2"
value="http://us.expasy.org/cgi-bin/niceprot.pl?$SEQUENCE_ID$">

                        </applet>
```

Appendix 19 – GeneMANIA

```
//where n is the current total of genes

For i= 1 to n

Executebutton.visible = True

If checkbox[i].item.selected > 3

      checkbox.visible = "Disable"

      clearbutton.visible = True

      Label[i].txt = checkbox[i]

Next i

Executebutton.click()

{

      iframe.visible = True

      Dim urlgenemania = "http://genemania.org&o=9606&g="

checkbox.select.count = j

For k=1 to j

message.txt = label(k) + '/'

End
```

This is translated into ASP.NET code as follows:

```
Private Sub Page_Load(ByVal sender As System.Object, ByVal e As
System.EventArgs) Handles MyBase.Load

            Dim index1 As Integer
            Label1.Text = "<p> Selected gene(s):</p>"
            Label2.Text = ""
            For index1 = 0 To CheckBoxList1.Items.Count - 1
                If (CheckBoxList1.Items(index1).Selected = "True") Then
                Label1.Text           =           Label1.Text           +
CheckBoxList1.Items(index1).Text + "   "
                    'If (CheckBoxList1.Items.Count > 5) Then
                    'CheckBoxList1.Visible = "False"
                    'End If
                Label2.Text= Label2.Text + CheckBoxList1.Items(index1).Text
+ "|"
                End If
            Next index1

            Dim         geneBaseUrl1          As          String          =
"http://genemania.org/link?o=9606&g="

            'Dim dv1 As DataView
            'Dim genedetails As String
            'Label2.Text = geneBaseUrl1 + genedetails
        HyperLink1.NavigateUrl = geneBaseUrl1 + Label2.Text
        HyperLink2.NavigateUrl = "~/overview/interaction.aspx"
```

```
        'HyperLink1.Target=
         'End If
    End Sub
```

The javascript used to embed the iframe code into the webpage is as shown:

```
<iframe                    name="Iframe1"                    runat="server"
style="width: 800px; height: 550px;">
<p>Your       browser       does       not       support       iframes.</p>
</iframe>
```

Appendix 20 – Google Earth API

```
<script
src="https://www.google.com/jsapi?key=ABQIAAAAJ_J8_jWNn_vcAAQ1M9fDLhTZZQSlEEZJ7DYzw
lZ4PnkrDTAOWBQIFqcf1aJ41x0lwJIR4jV7ENAhXA" type="text/javascript"></script>



<script type="text/javascript">
    var xmlDoc = null;
    if (window.ActiveXObject) {
// code for IE
        xmlDoc = new ActiveXObject("Microsoft.XMLDOM");
        xmlDoc.async = false;
        xmlDoc.load("countrylatlong.xml");
        var x = xmlDoc.getElementsByTagName('Alsod');
    }
    else if (document.implementation.createDocument) {
// code for Mozilla, Firefox, Opera, etc.
        xmlDoc = document.implementation.createDocument("", "", null);
        xmlDoc.async = false;
        xmlDoc.load("countrylatlong.xml");
    //xmlDoc.onload = createPlacemark1;
    //var x = xmlDoc.getElementsByTagName("Alsod");
        var x = xmlDoc.documentElement;

    }
    else {
        alert('Your browser cannot handle this script');
    }

        var ge = null;
        var placemark = new Array();
        var point = new Array();
        var la = new Array();
        var counter = 0;
```

```
        google.load("earth", "1");



    function init() {
        google.earth.createInstance('map3d', initCB,failureCB);
 // google.earth.createInstance("map3d", loadNetworkLink,failureCallback);
    }

    function initCB(instance) {
        ge = instance;
        ge.getWindow().setVisibility(true);  // required!
       ge.getNavigationControl().setVisibility(ge.VISIBILITY_AUTO);
       ge.getOptions().setOverviewMapVisibility(true);
       createPlacemark2();
    }



    function createPlacemark2() {



       var mutation = new Array();
       var latitude = new Array();
       var longitude = new Array();
       var country = new Array();

   // create arrays of placemarks,points etc

       for (i = 0; i < x.length; i++) {

           mutation[i]                                                                    =
(x[i].getElementsByTagName('mutation_mnemonic')[0].childNodes[0].nodeValue);
           latitude[i] = (x[i].getElementsByTagName('latitude')[0].childNodes[0].nodeValue);
```

```
        longitude[i] = (x[i].getElementsByTagName('longitude')[0].childNodes[0].nodeValue);
        country[i] = (x[i].getElementsByTagName('country')[0].childNodes[0].nodeValue);



        placemark[i] = ge.createPlacemark('');

// create the point
        point[i] = ge.createPoint('');
        point[i].setLatitude(latitude[i]);
        point[i].setLongitude(longitude[i]);
        placemark[i].setGeometry(point[i]);

// add the placemark to the earth DOM
        ge.getFeatures().appendChild(placemark[i]);

// look at the placemark we created
        la[i] = ge.createLookAt('');
        la[i].set(latitude[i], longitude[i],
          0, // altitude
          ge.ALTITUDE_RELATIVE_TO_GROUND,
          0, // heading
          0, // straight-down tilt
          4521188 // range (inverse of zoom)
          );

        ge.getView().setAbstractView(la[i]);

// give the placemark a name and a description (a balloon will
// automatically show on click)
        placemark[i].setName(mutation[i]);
        placemark[i].setDescription('Found in ' + country[i]);
      }
    }
```

```
    function failureCB(errorCode) {
        alert("Failure loading the Google Earth Plugin: " + errorCode);
    }

        google.setOnLoadCallback(init);
    </script>
<body onload="initialize()">
<div id="map3d_container" style="border: 1px solid silver; height: 500px;">
<div id="map3d" style="height: 100%;"></div>
</div>
</body>
```

Appendix 21 – Python Script for Population Frequency

```
#!/usr/bin/env python3


def is_pos(pos_str):
    ''' if the string is like '21:123455' or 'X:123' etc '''
    parts=pos_str.split(':')
    if len(parts) != 2:
        return False
    if parts[0] not in ('1','2','3','4','5','6','7','8','9','10','11','12','13', '14','15','16','17','18','19','20','21','22','X'):
        return False
    if parts[1].isdigit():
        return True
    return False

##############################################################################################
# read the list of SNP positions from the input table


snp_pos=set([])


ifile=open('population_frequency.txt')
for line in ifile:
    cols=line.split('\t')
    if is_pos(cols[6]):
        snp_pos.add(cols[6])
    else:
        continue


ifile.close()
# the input table is a bit missformatted on some lines (447,480)
# for example: line 480
# one extra tab after 'NP_001136.1:p.Phe100Ile'.
# The position string is now in column 8, instead of column 7.
# Removed it manually.
```

```
#################################################################################################
# for all files in directory 1kg
# read the SNP lines, ignore the indels.

import os
os.chdir('1kg')

snp_lines={}
for pos in snp_pos:
    snp_lines[pos]=''

for filename in os.listdir('.'):
    ifile=open(filename)
    for line in ifile:
        cols=line.split('\t')
        pos_str=cols[0]
        if pos_str in snp_pos:
            if len(cols[1])==1 and len(cols[2])==1: # SNP
                snp_lines[pos_str]=snp_lines[pos_str]+'\t'+line[:-1]
            else: # indels
                continue
    ifile.close()

os.chdir('..')

#################################################################################################
# re-open the input table, appending the lines from 1kg
ifile=open('population_frequency.txt')
for line in ifile:
    cols=line.split('\t')
    if cols[6] not in snp_pos:
        print(line[:-1])
```

```
    else:
        print(line[:-1]+'\t'+snp_lines[cols[6]])

ifile.close()
```

Appendix 22 – Perl Scripts to transpose, break, replace and join columns of a sequence

Transpose sequence

```
#!/usr/bin/perl
# Transpose lineS <-> 1 column
# OLUBUNMI ABEL
# 17/06/2011

$infile = "TDP43_codon.txt";
$outfile = "";

###  READ ARGUMENTS  ########################################################

foreach my $a (0..$#ARGV) {

  ### input directory
  if ($ARGV[$a] eq "-i") {
    $infile = $ARGV[$a+1];
  }

  ### output file
  elsif ($ARGV[$a] eq "-o") {
    $outfile = $ARGV[$a+1];
  }
  ### help
  elsif ($ARGV[$a] eq "-h") {
    die "Syntax: transpose_seq.pl -i inputfilename -o outputfilename (default output=inputfilename.tr)\n";
  }
}

if ($infile eq ""){die "STOP! You have to specify the input file name";}
```

```
if ($outfile eq ""){$outfile="$infile".".tr"}


###  READ AND WRITE DATA  ##########################################################

open inf, $infile or die "STOP! File $infile not found";
open ouf, ">$outfile";

$i=0;

foreach $line (<inf>){
 chomp $line;
  @line=split //,$line;
  for ($j=0;$j<=$#line;$j++){
    print ouf "$line[$j]\n";
  }
  $i++;
}

print "Output file $outfile created\n";

close inf;
close ouf;
```

Join columns

```
#!/usr/bin/perl
#
# Join-cols
#
# Olubunmi Abel
# 05/06/2009

&ReadArguments;
```

```
&CreateNewTable;

################################################
### Read arguments from the command line

sub ReadArguments {

$verbo=0;
$infile1="SOD1_sequence_twenties.txt.tr";
$infile2="SOD1_codon.txt.tr";
$shift=0;
$outfile="result1.txt";
$th1="";
$th2="";

foreach my $a (0..$#ARGV) {

        ### help
        if ($ARGV[0] eq "-h") {
                &PrintHelp;
    }

        ### input file 1
        elsif ($ARGV[$a] eq "-i1") {
                $infile1 = $ARGV[$a+1];
        }

        ### input file 2
        elsif ($ARGV[$a] eq "-i2") {
                $infile2 = $ARGV[$a+1];
        }
```

```
        ### shift
        elsif ($ARGV[$a] eq "-shift") {
                $shift = $ARGV[$a+1];
        }

        ### threshold
        elsif ($ARGV[$a] eq "-threshold") {
          if ($ARGV[$a+1] =~ /^-?\d+\.?\d*$/) {
            $th1 = $ARGV[$a+1];
            if ($ARGV[$a+2] =~ /^-?\d+\.?\d*$/) {
              $th2=$ARGV[$a+2];
         }
            else {$th2=$th1;
            }
          }
        }

        ### output file
        elsif ($ARGV[$a] eq "-o") {
                $outfile = $ARGV[$a+1];
        }

        ### verbosity
        elsif ($ARGV[$a] eq "-v") {
                $verbo=1;
    }
}

if ($infile1 eq "" or $infile2 eq "" ) {
        die "STOP! You have to give the name of to input file.\n";
}
if ($outfile eq "") {
        $outfile="$infile1_$infile2.out";
```

```
}

}  # End of ReadArguments

#########################################################################
### Print help



sub PrintHelp {
  open HELP, "| more";
  print <<EndHelp;
NAME
       join-cols.pl

DESCRIPTION
       Build one table (in one output file) from 2 separated
       tables (in 2 separated files).

AUTHORS
       Didier Gonze (dgonze\@ulb.ac.be)

OPTIONS
       -i1 in_file_name_1
       -i2 in_file_name_2
              Specify the input files name.
                     Two files must be specified.

      -shift #
            Specify the shift between file1 and file2 (default: shift=0).
                The shift can be negative.

           -threshold # [#]
                     Define a threshold below which the data are not printed.
```

```
                If only one value is given, the threshold will be the same for
                both files. If two values are given, the first one corresponds
                to the first file, while the second corresponds to the second.

        -o out_file_name
                Specify the output file. If no data in file 2 correspond
                to data in file 1 or vice versa (because of shift for example),
                        "NA" will be indicated.

        -v
                Verbosity: print detailed informations during the process.

        -h
                Give help (print this message). This argument must be
                the first one.

EXAMPLE
        perl join-cols.pl -i1 mylist1 -i2 mylist2 -shift 2 -threshold 0 -o mydoublelist

EndHelp
  close HELP;
die "\n";
}


#####################################################################
### Create the new table joining the two tables

sub CreateNewTable {

$nblines=0;
$nbnull=0;
```

```
open inf1, $infile1 or die "STOP! File $infile1 not found";
open inf2, $infile2 or die "STOP! File $infile2 not found";
open ouf, ">$outfile";

if ($shift > 0 ){
  for ($i=1; $i<=$shift; $i++){
    $line2=<inf2>;   # skip lines in inf2
    chomp $line2;
    if ("$th2" eq "" or $line2 >= $th2){
       print ouf "NA\t$line2\n";
    }
    $nbnull++;
  }
}
if ($shift < 0 ){
  for ($i=1; $i<=-$shift; $i++){
    $line1=<inf1>;   # skip lines in inf1
    chomp $line1;
    if ("$th1" eq "" or $line1 >= $th1){
       print ouf "$line1\tNA\n";
    }
    $nbnull++;
  }
}

foreach $line1 (<inf1>) {
    $nblines++;
    chomp $line1;
    $line2=<inf2>;
    if ($line2 ne ""){
      chomp $line2;
      if (("$th1" eq "" or $line1 >= $th1) and ("$th2" eq "" or $line2 >= $th2)){
         print ouf "$line1\t$line2\n";
```

```
      }
    }
    else{
      if ("$th1" eq "" or $line1 >= $th1) {
         print ouf "$line1\tNA\n";  # add last lines from inf1
      }
      $nbnull++;
    }
}

if ($shift < 0 ){
  for ($i=1; $i<=-$shift; $i++){
    $line2=<inf2>;   # add last lines from inf2
    chomp $line2;
    if ("$th2" eq "" or $line2 >= $th2){
       print ouf "NA\t$line2\n";
    }
    $nbnull++;
  }
}

print "Output file $outfile created\n";
close inf1;
close inf2;
close ouf;

if ($verbo==1) {print "File $outfile created.\n";}
if ($verbo==1) {print "There are $nblines OK lines.\n";}
if ($verbo==1) {print "There are $nbnull NULL lines.\n";}
if ($verbo==1) {print "Shift=$shift.\n";}

}  # End of CreateNewTable
```

Repeat amino acids 20 times

```
#!/usr/bin/perl


# Transpose lineS <-> 1 column
# oLUBUNMI ABEL
# 05/06/2009


$outfile = "DCTN1_possibilities.txt";



###  READ AND WRITE DATA  ##################################################


open ouf, ">$outfile";


#### Let i be numberof codons (with M is different from without M)
for ($i = 1; $i <=1253; ($i=$i+1)) {
$j="A   R       N       D       C       Q       E       G       H       I       L       K       M
        F       P       S       T       W       Y       V";


print ouf $j;
print ouf "\n";


}
print "Output file $outfile created\n";
close ouf;
```

Break and replace columns

```
#!/usr/bin/perl


# READ COLUMNS, REPLACE SPACE WITH TAB AND ADD PRINT
# Olubunmi Abel
# 12/07/2011
```

```
$infile = "VCP_NETPHOS.txt";
$outfile = "";
###  READ ARGUMENTS  ######################################################
foreach my $a (0..$#ARGV) {
  ### input directory
  if ($ARGV[$a] eq "-i") {
    $infile = $ARGV[$a+1];
  }
  ### output file
  elsif ($ARGV[$a] eq "-o") {
    $outfile = $ARGV[$a+1];
  }
  ### help
  elsif ($ARGV[$a] eq "-h") {
    die "Syntax: transpose.pl -i inputfilename -o outputfilename (default output=inputfilename.tr)\n";
  }
}
if ($infile eq ""){die "STOP! You have to specify the input file name";
}
if ($outfile eq ""){$outfile="file_"."$infile"}
###  READ DATA  ###########################################################
open inf, $infile or die "STOP! File $infile not found";
open ouf, ">$outfile";



foreach $line (<inf>){
chomp $line;
@data=split / /,$line;

print ouf "VCP"."\t".@data[1]."\t".@data[2]."\t".@data[3]."\t".@data[4]."\n";
}print "Output file $outfile created\n";
close inf;
close ouf;
```

Breaking columns

```
#!/usr/bin/perl


# READ COLUMNS, REPLACE SPACE WITH TAB AND ADD PRINT
# Olubunmi Abel
# 12/07/2011
$infile = "VCP_NETPHOS.txt";
$outfile = "";
###  READ ARGUMENTS  ######################################################
foreach my $a (0..$#ARGV) {
  ### input directory
  if ($ARGV[$a] eq "-i") {
    $infile = $ARGV[$a+1];
  }
  ### output file
  elsif ($ARGV[$a] eq "-o") {
    $outfile = $ARGV[$a+1];
  }
  ### help
  elsif ($ARGV[$a] eq "-h") {
    die "Syntax: transpose.pl -i inputfilename -o outputfilename (default output=inputfilename.tr)\n";
  }
}
if ($infile eq ""){die "STOP! You have to specify the input file name";
}
if ($outfile eq ""){$outfile="file_"."$infile"}
###  READ DATA  ###########################################################
open inf, $infile or die "STOP! File $infile not found";
open ouf, ">$outfile";



foreach $line (<inf>){
```

```
chomp $line;
@data=split / /,$line;

print ouf "VCP"."\t".@data[1]."\t".@data[2]."\t".@data[3]."\t".@data[4]."\n";
}print "Output file $outfile created\n";
close inf;
close ouf;
```

Appendix 23 – SQL Script for testing and verifying ranking system

```
SELECT Rank_Mutations, Rank_Patients, Rank_Cases, Rank_Controls,
Rank_Codon, Rank_FALS, Rank_SALS, Rank_Replications, Rank_Pathogenicity,
Rank_Populations, Gene ,SUM( Rank_Mutations + Rank_Patients + Rank_Cases +
Rank_Controls + Rank_Codon + Rank_FALS + Rank_SALS + Rank_Replications +
Rank_Pathogenicity + Rank_Populations + 0) As Rank_Sum, DENSE_RANK() OVER
(ORDER BY SUM( Rank_Mutations + Rank_Patients + Rank_Cases + Rank_Controls
+ Rank_Codon + Rank_FALS + Rank_SALS + Rank_Replications +
Rank_Pathogenicity + Rank_Populations + 0) ASC) AS Final_Rank FROM (SELECT
[KCLAD\spngoka].Number_of_mutations_per_gene.Rank As
Rank_Mutations,[KCLAD\spngoka].Number_of_affected_patients_in_ALSoD.Rank AS
Rank_Patients, [KCLAD\spngoka].[Number_of_cases_recorded].Rank As
Rank_Cases, [KCLAD\spngoka].[Number_of_controls_recorded].Rank As
Rank_Controls, [KCLAD\spngoka].[Number_of_mutations_in
_same_codon_by_rank].Rank As Rank_Codon,
[KCLAD\spngoka].[Number_of_patients_with_family_history_FALS].Rank As
Rank_FALS,
[KCLAD\spngoka].[Number_of_patients_without_family_history_SALS].Rank As
Rank_SALS, [KCLAD\spngoka].[Number_of_times_mutation_is_replicated].Rank As
Rank_Replications,
[KCLAD\spngoka].[Number_of_pathogenic_mutations_by_rank].Rank As
Rank_Pathogenicity,
[KCLAD\spngoka].[Number_of_unique_countries_on_genes].Rank As
Rank_Populations, [KCLAD\spngoka].[Number_of_mutations_per_gene].Gene FROM
[KCLAD\spngoka].Number_of_affected_patients_in_ALSoD INNER JOIN
[KCLAD\spngoka].Number_of_mutations_per_gene ON
[KCLAD\spngoka].Number_of_affected_patients_in_ALSoD.Gene =
[KCLAD\spngoka].Number_of_mutations_per_gene.Gene INNER JOIN
[KCLAD\spngoka].[Number_of_cases_recorded] ON
[KCLAD\spngoka].[Number_of_mutations_per_gene].Gene =
[KCLAD\spngoka].[Number_of_cases_recorded].Gene INNER JOIN
[KCLAD\spngoka].[Number_of_controls_recorded] ON
[KCLAD\spngoka].[Number_of_mutations_per_gene].Gene =
[KCLAD\spngoka].[Number_of_controls_recorded].Gene INNER JOIN
[KCLAD\spngoka].[Number_of_mutations_in _same_codon_by_rank] ON
[KCLAD\spngoka].[Number_of_mutations_per_gene].Gene =
[KCLAD\spngoka].[Number_of_mutations_in _same_codon_by_rank].Gene INNER
JOIN [KCLAD\spngoka].[Number_of_patients_with_family_history_FALS] ON
[KCLAD\spngoka].[Number_of_mutations_per_gene].Gene =
[KCLAD\spngoka].[Number_of_patients_with_family_history_FALS].Gene INNER
JOIN [KCLAD\spngoka].[Number_of_patients_without_family_history_SALS] ON
[KCLAD\spngoka].[Number_of_mutations_per_gene].Gene =
[KCLAD\spngoka].[Number_of_patients_without_family_history_SALS].Gene INNER
JOIN [KCLAD\spngoka].[Number_of_times_mutation_is_replicated] ON
[KCLAD\spngoka].[Number_of_mutations_per_gene].Gene =
[KCLAD\spngoka].[Number_of_times_mutation_is_replicated].Gene INNER JOIN
[KCLAD\spngoka].[Number_of_pathogenic_mutations_by_rank] ON
[KCLAD\spngoka].[Number_of_mutations_per_gene].Gene =
[KCLAD\spngoka].[Number_of_pathogenic_mutations_by_rank].Gene INNER JOIN
[KCLAD\spngoka].[Number_of_unique_countries_on_genes] ON
```

```
[KCLAD\spngoka].[Number_of_mutations_per_gene].Gene =
[KCLAD\spngoka].[Number_of_unique_countries_on_genes].Gene) As table1 GROUP
BY Rank_Mutations, Rank_Patients, Rank_Cases, Rank_Controls, Rank_Codon,
Rank_FALS, Rank_SALS, Rank_Replications, Rank_Pathogenicity,
Rank_Populations, Gene
```

Appendix 24 – Columns/Field list in each SQL table

Gene

```
SELECT [gene_id]
      ,[hgnc_id]
      ,[ensembl_id]
      ,[swissport_id]
      ,[ncbi_locuslink_id]
      ,[ncbi_refseq_id]
      ,[structure_id]
      ,[omim_id]
      ,[genecards_id]
      ,[gene_name]
      ,[keywords]
      ,[chromosome_name]
      ,[chromosome_position]
      ,[chromosome_band]
      ,[chromosome_bp]
      ,[chromosome]
      ,[protein_name]
      ,[protein_function]
      ,[phenotype]
      ,[gene_certification]
      ,[gene_comment]
      ,[other_names]
      ,[sequence_tran]
      ,[sequence]
      ,[ucsc_id]
      ,[accession_no]
      ,[reason_for_investigation]
      ,[result]
      ,[category]
      ,[graph]
      ,[gene_effect]
      ,[ihop]
      ,[pdb_id]
      ,[dbSNP]
      ,[pubmed_id1]
      ,[pubmed_id2]
      ,[ALSgene_id]
      ,[inheritance_pattern]
      ,[syndrome_id]
      ,[causative_id]
      ,[FALS_only]
      ,[DB_no]
      ,[snp]
      ,[basepair]
      ,[pvalue]
      ,[snp_pubmed_id]
      ,[snp_paperlink]
      ,[lifesciencedb]
  FROM [Alsod].[dbo].[gene]
```

Mutation

```
SELECT [mutation_id]
      ,[mutation_alias]
      ,[mutation_name]
      ,[mutation_mnemonic]
      ,[gene_id]
      ,[institution_code]
      ,[mutation_type]
```

```
      ,[seq_location_id]
      ,[position]
      ,[sequence_position_relative]
      ,[codon]
      ,[codon_id]
      ,[sequence_original]
      ,[sequence_mutated]
      ,[sequence_location_type]
      ,[sequence_location_number]
      ,[aa_original]
      ,[aa_mutated]
      ,[restriction_site]
      ,[enzyme]
      ,[mutation_approved_status]
      ,[mutation_approved_date]
      ,[mutation_approved_by]
      ,[mutation_submitted_date]
      ,[mutation_documentation]
      ,[mutation_comment]
      ,[first_author]
      ,[year]
      ,[journal]
      ,[mutation_documentation_type]
      ,[link]
      ,[phenotype]
      ,[zygosity]
      ,[reference_id]
      ,[countryiso1]
      ,[countryiso2]
      ,[countryiso3]
      ,[countryiso4]
      ,[countryiso5]
      ,[countryiso6]
      ,[countryiso7]
      ,[countryiso8]
      ,[countryiso9]
      ,[countryiso10]
      ,[countryiso11]
      ,[countryiso12]
      ,[countryiso13]
      ,[countryiso14]
      ,[countryiso15]
      ,[countryiso16]
      ,[countryiso17]
      ,[countryiso18]
      ,[countryiso19]
      ,[countryiso20]
      ,[swissprot_id]
      ,[pubmed_id]
      ,[title]
      ,[doi]
      ,[dbSNP]
      ,[frequency]
      ,[frequency_references]
      ,[data_from]
      ,[HGVS_Nucleotide]
      ,[HGVS_protein]
      ,[Location]
      ,[dbSNP_id]
  FROM [Alsod].[dbo].[mutation]
```

Patient_clinical_record

```
SELECT [patient_id]
      ,[sensory_symptom]
      ,[sensory_sign]
      ,[autonomic_dysfunction_symptom]
      ,[autonomic_dysfunction_sign]
      ,[parkinsonism_symptom]
      ,[parkinsonism_sign]
      ,[dementia]
      ,[dementia_description]
      ,[other_atypical_exists]
      ,[other_atypical_description]
      ,[neuropathology_available]
      ,[s_path_mcx]
      ,[s_path_cort_spin_tract]
      ,[s_path_ahmn]
      ,[s_path_post_column]
      ,[s_path_spin_cerebell_tra]
      ,[s_path_oculomot_nucl]
      ,[s_path_onuf_nucl]
      ,[s_path_clarke_column]
      ,[s_path_other_description]
      ,[m_path_mcx_bunina_body]
      ,[m_path_mcx_LBI]
      ,[m_path_mcx_UbIRI]
      ,[m_path_mcx_vacuola]
      ,[m_path_mcx_nf_acc_inn_p]
      ,[m_path_mcx_nf_acc_inn_nonp]
      ,[m_path_mcx_nf_acc_axon]
      ,[m_path_mcx_sod1_IR]
      ,[m_path_sc_bunina_body]
      ,[m_path_sc_LBI]
      ,[m_path_sc_UbIRI]
      ,[m_path_sc_vacuola]
      ,[m_path_sc_nf_acc_inn_p]
      ,[m_path_sc_nf_acc_inn_nonp]
      ,[m_path_sc_nf_acc_axon]
      ,[m_path_sc_sod1_ir]
      ,[m_path_prfront_bunina_body]
      ,[m_path_prfront_lbi]
      ,[m_path_prfront_ubiri]
      ,[m_path_prfront_vacuola]
      ,[m_path_prfront_nf_acc_inn_p]
      ,[m_path_prfront_nf_acc_inn_nonp]
      ,[m_path_prfront_nf_acc_axon]
      ,[m_path_prfront_sod1_ir]
      ,[m_path_hippocamp_ubiri]
      ,[s_path_other_desc_short]
      ,[genetics_pub_exists]
      ,[genetics_pub_desc_short]
      ,[genetics_pub_desc_long]
      ,[clinical_pub_exists]
      ,[clinical_pub_desc_long]
      ,[clinical_pub_desc_short]
      ,[neuropath_pub_exists]
      ,[neuropath_pub_desc_long]
      ,[neuropath_pub_desc_short]
      ,[experimental_pub_exists]
      ,[experimental_pub_desc_long]
      ,[experimental_pub_desc_short]
      ,[activity_rbc_method_id]
      ,[activity_rbc_method]
      ,[activity_rbc_value]
```

```
,[activity_lymphoblast_method_id]
,[activity_lymphoblast_method]
,[activity_lymphoblast_value]
,[activity_fibroblast_method_id]
,[activity_fibroblast_method]
,[activity_fibroblast_value]
,[activity_brain_method_id]
,[activity_brain_method]
,[activity_brain_value]
,[activity_cervical_method_id]
,[activity_cervical_method]
,[activity_cervical_value]
,[activity_thoracic_method_id]
,[activity_thoracic_method]
,[activity_thoracic_value]
,[activity_lumbar_method_id]
,[activity_lumbar_method]
,[activity_lumbar_value]
,[activity_other_method_id]
,[activity_other_method]
,[activity_other_method_value]
,[presenting_symptom]
,[cra_emotional]
,[cra_jaw_jerk]
,[cra_facial_reflexes]
,[cra_tongue_fasc]
,[cra_tongue_wast]
,[cra_tongue_spast]
,[ul_left_wasting]
,[ul_left_fasciculation]
,[ul_left_weakness]
,[ul_left_tone]
,[ul_left_triceps_reflex]
,[ul_left_biceps_reflex]
,[ul_left_brachiorad_reflex]
,[ul_left_fingerHS]
,[ul_right_wasting]
,[ul_right_fasciculation]
,[ul_right_weakness]
,[ul_right_tone]
,[ul_right_triceps_reflex]
,[ul_right_biceps_reflex]
,[ul_right_brachiorad_reflex]
,[ul_right_fingerHS]
,[thor_trunk_wasting]
,[thor_trunk_fasciculation]
,[thor_trunk_weakness]
,[thor_trunk_tone]
,[ll_left_wasting]
,[ll_left_fasciculation]
,[ll_left_weakness]
,[ll_left_tone]
,[ll_left_clonus]
,[ll_left_knee_reflex]
,[ll_left_ankle_reflex]
,[ll_left_plantar]
,[ll_right_wasting]
,[ll_right_fasciculation]
,[ll_right_weakness]
,[ll_right_tone]
,[ll_right_clonus]
,[ll_right_knee_reflex]
```

```
      ,[ll_right_ankle_reflex]
      ,[ll_right_plantar]
      ,[umn_rostral_lmn]
      ,[mr_imaging_head]
      ,[mr_imaging_cerv_spine]
      ,[mr_imaging_tl_spine]
      ,[evid_lmn_involve_cra_emg]
      ,[evid_lmn_involve_ul_left_emg]
      ,[evid_lmn_involve_ul_right_emg]
      ,[evid_lmn_involve_thortrunk_emg]
      ,[evid_lmn_involve_ll_left_emg]
      ,[evid_lmn_involve_ll_right_emg]
      ,[if_umn_abs_were_they_ever_pr]
      ,[el_escorial_criteria]
      ,[pat_nonneuro_disease_descr]
      ,[fam_nonneuro_disease_exists]
      ,[fam_nonneuro_disease_descr]
      ,[pat_nonneuro_disease_short]
      ,[fam_nonneuro_disease_short]
      ,[material_dna]
      ,[material_lymphoblast]
      ,[material_fibroblast]
      ,[material_brain]
      ,[material_spinal_cord_cervical]
      ,[material_spinal_cord_thoracic]
      ,[material_spinal_cord_lumbar]
      ,[material_other]
  FROM [Alsod].[dbo].[patient_clinical_record]
```

Patient_genetic_record

```
SELECT [patient_id]
      ,[gene_id]
      ,[family_id]
      ,[screened]
      ,[mutation_id]
      ,[mutation_found]
      ,[zygosity]
      ,[status]
      ,[family_history]
      ,[affected]
  FROM [Alsod].[dbo].[patient_genetic_record]
```

Patient_record

```
SELECT [patient_id]
      ,[institution_code]
      ,[patient_mnemonic]
      ,[patient_first_name]
      ,[patient_middle_init]
      ,[patient_last_name]
      ,[family_id]
      ,[father_id]
      ,[mother_id]
      ,[family_history_old]
      ,[sex]
      ,[birth_date]
      ,[country_iso_code]
      ,[ethnic_origin]
      ,[status_old]
      ,[screened_for_sod1_old]
      ,[mutation_found_old]
      ,[mutation_id_old]
```

```
      ,[zygosity]
      ,[last_accessed]
      ,[date_of_onset]
      ,[date_of_death]
      ,[date_of_diagnosis]
      ,[alive]
      ,[age_of_onset]
      ,[diagnosis_age]
      ,[death_age]
      ,[site_of_onset]
      ,[site_of_onset_lr]
      ,[site_of_onset_pd]
      ,[death_diagnosis]
      ,[weight]
      ,[fvc]
      ,[disease_duration]
      ,[comments]
      ,[UMN_signs]
      ,[LMN_signs]
      ,[cognitive_signs]
      ,[phenotype]
  FROM [Alsod].[dbo].[patient_record]
```

Replicated mutation table

```
SELECT [gene]
      ,[mutation]
      ,[codon]
      ,[exons]
      ,[families]
      ,[generations]
      ,[affected_patients]
      ,[affected_controls]
      ,[sporadic]
      ,[familial]
      ,[family_history]
      ,[cases]
      ,[controls]
      ,[LODscore]
      ,[linkage_shown]
      ,[country1]
      ,[country2]
      ,[country3]
      ,[country4]
      ,[country5]
      ,[ethnicity]
      ,[snp]
      ,[pathogenic]
      ,[author]
      ,[year]
      ,[pubmed_id]
      ,[PANTHER_Value]
      ,[PANTHER_Prediction]
      ,[SIFT_Value]
      ,[SIFT_Prediction]
      ,[PolyPhen_Value]
      ,[PolyPhen_Prediction]
      ,[controversies]
      ,[comment]
      ,[S_No]
      ,[title]
      ,[link]
      ,[country6]
```

```
      ,[country7]
      ,[country8]
      ,[country9]
      ,[country10]
  FROM [Alsod].[dbo].[mutation_replication]
```

Animal models

```
SELECT [HomoloGeneID]
      ,[Organism_Name]
      ,[Biological_Name]
      ,[Kegg_id]
      ,[ncbi_locuslink_id]
      ,[gene_id]
      ,[EntrezGene_id]
      ,[MGI_id]
      ,[hgnc_id]
      ,[omim_id]
      ,[chromosome]
      ,[genomic_coordinates_strand]
      ,[genomic_chromosome]
      ,[genomic_location]
      ,[chromosome_split]
      ,[genomic_range]
      ,[strand]
      ,[gene_name]
      ,[Synonyms]
      ,[ID]
  FROM [Alsod].[dbo].[AnimalModels]
```

Codons

```
SELECT [gene_id]
      ,[trinucleotide]
      ,[codon_id]
      ,[sequence_position_list]
      ,[sequence_location_number]
  FROM [Alsod].[dbo].[codons]
```

Continents

```
SELECT [continent_id]
      ,[continent_name]
  FROM [Alsod].[dbo].[Continent]
```

Counter

```
SELECT [count]
  FROM [Alsod].[dbo].[counter]
```

Country

```
SELECT [country_iso_code]
      ,[country_name]
      ,[continent_id]
  FROM [Alsod].[dbo].[Country]
```

Country details

```
SELECT [country_iso_code]
      ,[country]
      ,[fips104]
      ,[iso2]
```

```
      ,[iso3]
      ,[isono]
      ,[capital]
      ,[region]
      ,[currency]
      ,[currency_code]
      ,[population]
      ,[t_lag]
      ,[t_type]
      ,[comments]
      ,[language]
      ,[city_info]
      ,[dialing_code]
      ,[latitude]
      ,[longitude]
  FROM [Alsod].[dbo].[countrydetails]
```

Country information

```
SELECT [country_iso_code]
      ,[country]
      ,[fips104]
      ,[iso2]
      ,[iso3]
      ,[isono]
      ,[capital]
      ,[region]
      ,[currency]
      ,[currency_code]
      ,[population]
      ,[t_lag]
      ,[t_type]
      ,[comments]
      ,[language]
      ,[city_info]
      ,[dialing_code]
  FROM [Alsod].[dbo].[countryinfo]
```

Country latitude and longitude

```
SELECT [country_iso_code]
      ,[latitude]
      ,[longitude]
  FROM [Alsod].[dbo].[countrylatlong]
```

Credibility_survey

```
SELECT [S_No]
      ,[Expert]
      ,[ALS2]
      ,[ANG]
      ,[DAO]
      ,[DCTN1]
      ,[FIG4]
      ,[FUS]
      ,[NEFH]
      ,[OPTN]
      ,[SETX]
      ,[SOD1]
      ,[SPG11]
      ,[TARDBP]
      ,[VAPB]
      ,[VCP]
```

```
  FROM [Alsod].[dbo].[credibility_survey]
```

Gene authors

```
SELECT [AuthorID]
      ,[Lastname]
      ,[Firstname]
      ,[Initials]
      ,[Address]
      ,[Email]
  FROM [Alsod].[dbo].[gene_authors]
```

Gene study

```
SELECT [Lastname]
      ,[AuthorID]
      ,[Month]
      ,[Year]
      ,[Others]
      ,[Title]
      ,[Volume]
      ,[Paperlink]
      ,[gene_id]
      ,[PMID]
      ,[Comment]
      ,[doi_key]
      ,[StudyID]
  FROM [Alsod].[dbo].[gene_study]
```

Gene_sequence

```
SELECT [gene_id]
      ,[sequence_position]
      ,[sequence_position_relative]
      ,[sequence_location_type]
      ,[sequence_location_number]
      ,[sequence_location_order]
      ,[nucleotide]
  FROM [Alsod].[dbo].[gene_sequence]
```

Epidemiology

```
SELECT [Gene]
      ,[country]
      ,[country_iso_code]
      ,[set]
      ,[controls]
      ,[cases]
      ,[Male-cases]
      ,[Female-cases]
      ,[bulbar]
      ,[limb]
      ,[others]
      ,[dead]
      ,[alive]
      ,[unknown]
      ,[FALS]
      ,[SALS]
      ,[first_author]
      ,[year]
      ,[title]
      ,[pubmed_id]
      ,[phenotype]
```

```
  FROM [Alsod].[dbo].[Epidemiology]
```

Gene frequency

```
SELECT [s_no]
      ,[gene]
      ,[country_iso_code]
      ,[replication_id]
      ,[Total_cases]
      ,[Total_FALS_screened]
      ,[Total_SALS_screened]
      ,[Total_affected]
      ,[FALS_number]
      ,[FALS_percentage]
      ,[SALS_number]
      ,[SALS_percentage]
      ,[Total_controls]
      ,[Total_controls_affected]
      ,[phenotype]
      ,[comment]
  FROM [Alsod].[dbo].[gene_frequency]
```

Gene snps

```
SELECT [gene_id]
      ,[snp]
      ,[basepair]
      ,[pvalue]
      ,[pubmed_id]
      ,[paperlink]
      ,[term]
      ,[S_No]
      ,[First_Author]
      ,[Year]
  FROM [Alsod].[dbo].[gene_SNP]
```

Gene snps ALSGene website

```
SELECT [chromosome_name]
      ,[gene_id]
      ,[ncbi_locuslink_id]
      ,[gene_name]
      ,[First_Author]
      ,[Year]
      ,[pubmed_id]
      ,[snp1]
      ,[snp2]
      ,[mutation_mnemonic]
      ,[codon]
      ,[paperlink]
      ,[basepair]
      ,[term]
      ,[pvalue]
      ,[S_No]
  FROM [Alsod].[dbo].[gene_SNP_ALSGene]
```

Gene snps Lifeseq website

```
SELECT [gene_id]
      ,[DNA_change_original_from]
      ,[DNA_change_original_to]
      ,[chromosome_name]
      ,[position]
```

```
      ,[basepair]
      ,[exon]
      ,[snp2]
      ,[mutation_mnemonic]
      ,[codon]
      ,[mRNA_accession_no]
      ,[amino_acid_accession_no]
      ,[amino_acid_change_as_seen]
      ,[zygosity]
      ,[ethnicity]
      ,[country]
      ,[families]
      ,[number_patients_with_mutations]
      ,[rate_patients_with_mutation]
      ,[rate_patients]
      ,[rate_patients_mutation_found]
      ,[rate_patients_mutation_not_found]
      ,[rate_controls]
      ,[rate_controls_mutation_found]
      ,[rate_controls_mutation_not found]
      ,[odd_ratio]
      ,[CI(95%)]
      ,[Chi-Square_p_value]
      ,[clinical_characteristics]
      ,[family_history]
      ,[gender]
      ,[age_onset]
      ,[duration_in_text]
      ,[duration_months]
      ,[UMN]
      ,[LMN]
      ,[Site_onset]
      ,[years_until_initiation_respirator]
      ,[pubmed_id]
      ,[paper_title]
      ,[First_Author]
      ,[Year]
      ,[comment]
      ,[journal]
  FROM [Alsod].[dbo].[gene_SNP_lifeseq]
```

Gene snps updated

```
SELECT [gene_id]
      ,[snp]
      ,[basepair]
      ,[pvalue]
      ,[pubmed_id]
      ,[paperlink]
      ,[term]
      ,[First_Author]
      ,[Year]
      ,[link]
      ,[Amino_acid]
      ,[S_No]
  FROM [Alsod].[dbo].[gene_SNP_updated]
```

Guestbook

```
SELECT [LastName]
      ,[FirstName]
      ,[Email]
      ,[Comment]
```

```
      ,[SentDateTime]
  FROM [Alsod].[dbo].[guestbook]
```

GWA_BOS

```
SELECT [CHROMOSOME]
      ,[SNP]
      ,[BP]
      ,[A1]
      ,[F_A]
      ,[F_U]
      ,[A2]
      ,[CHISQ]
      ,[PVALUE]
      ,[ODDRATIO]
      ,[L95]
      ,[U95]
      ,[neglogpval]
  FROM [Alsod].[dbo].[GWA_BOS]
```

GWA_FRA

```
SELECT [CHROMOSOME]
      ,[SNP]
      ,[BP]
      ,[A1]
      ,[F_A]
      ,[F_U]
      ,[A2]
      ,[CHISQ]
      ,[PVALUE]
      ,[ODDRATIO]
      ,[L95]
      ,[U95]
      ,[neglogpval]
  FROM [Alsod].[dbo].[GWA_FRA]
```

GWA_HOL

```
SELECT [CHROMOSOME]
      ,[SNP]
      ,[BP]
      ,[A1]
      ,[F_A]
      ,[F_U]
      ,[A2]
      ,[CHISQ]
      ,[PVALUE]
      ,[ODDRATIO]
      ,[L95]
      ,[U95]
      ,[neglogpval]
  FROM [Alsod].[dbo].[GWA_HOL]
```

GWA_NIH

```
SELECT [CHROMOSOME]
      ,[SNP]
      ,[BP]
      ,[A1]
      ,[F_A]
      ,[F_U]
      ,[A2]
```

```
      ,[CHISQ]
      ,[PVALUE]
      ,[ODDRATIO]
      ,[L95]
      ,[U95]
      ,[neglogpval]
  FROM [Alsod].[dbo].[GWA_NIH]
```

GWA_NOKEY

```
SELECT [CHROMOSOME]
      ,[SNP]
      ,[BP]
      ,[A1]
      ,[F_A]
      ,[F_U]
      ,[A2]
      ,[CHISQ]
      ,[PVALUE]
      ,[ODDRATIO]
      ,[L95]
      ,[U95]
      ,[ID]
      ,[POP]
  FROM [Alsod].[dbo].[GWA_NOKEY]
```

GWA_UK

```
SELECT [CHROMOSOME]
      ,[SNP]
      ,[BP]
      ,[A1]
      ,[F_A]
      ,[F_U]
      ,[A2]
      ,[CHISQ]
      ,[PVALUE]
      ,[ODDRATIO]
      ,[L95]
      ,[U95]
      ,[neglogpval]
  FROM [Alsod].[dbo].[GWA_UK]
```

GWA_USA

```
SELECT [CHROMOSOME]
      ,[SNP]
      ,[BP]
      ,[A1]
      ,[F_A]
      ,[F_U]
      ,[A2]
      ,[CHISQ]
      ,[PVALUE]
      ,[ODDRATIO]
      ,[L95]
      ,[U95]
      ,[neglogpval]
  FROM [Alsod].[dbo].[GWA_USA]
```

GWA_UK_BOS

```
SELECT [chr]
```

```
      ,[snp]
      ,[bp]
      ,[pvalbos]
      ,[pvaluk]
      ,[newpvalueukbos]
      ,[id]
      ,[log_ukbos]
  FROM [Alsod].[dbo].[GWA_UK_BOS]
```

GWA_UK_BOS_FRA

```
SELECT [chr]
      ,[snp]
      ,[bp]
      ,[pvalfra]
      ,[pvalbos]
      ,[newpvaluebosfra]
      ,[pvaluk]
      ,[newpvalueukbosfra]
      ,[id]
      ,[log_bosfra]
      ,[log_ukbosfra]
  FROM [Alsod].[dbo].[GWA_UK_BOS_FRA]
```

GWA_UK_BOS_FRA_NIH

```
SELECT [chr]
      ,[snp]
      ,[bp]
      ,[pvalusa]
      ,[pvalfra]
      ,[newpvaluefrausa]
      ,[pvalbos]
      ,[newpvaluebosfrausa]
      ,[pvaluk]
      ,[newpvalueukbosfrausa]
      ,[newpvalueukfrausa]
      ,[newpvalueukfra]
      ,[newpvaluebosusa]
      ,[newpvalueukusa]
      ,[id]
      ,[log_frausa]
      ,[log_bosfrausa]
      ,[log_ukbosfrausa]
      ,[log_ukfrausa]
      ,[log_ukfra]
      ,[log_bosusa]
      ,[log_ukusa]
  FROM [Alsod].[dbo].[GWA_UK_BOS_FRA_NIH]
```

GWA_UK_BOS_FRA_USA

```
SELECT [chr]
      ,[snp]
      ,[bp]
      ,[pvalusa]
      ,[pvalfra]
      ,[newpvaluefrausa]
      ,[pvalbos]
      ,[newpvaluebosfrausa]
      ,[pvaluk]
      ,[newpvalueukbosfrausa]
      ,[newpvalueukfrausa]
```

```
      ,[newpvalueukfra]
      ,[newpvaluebosusa]
      ,[newpvalueukusa]
      ,[id]
      ,[log_frausa]
      ,[log_bosfrausa]
      ,[log_ukbosfrausa]
      ,[log_ukfrausa]
      ,[log_ukfra]
      ,[log_bosusa]
      ,[log_ukusa]
  FROM [Alsod].[dbo].[GWA_UK_BOS_FRA_USA]
```

GWA_UK_BOS_HOL_FRA_NIH

```
SELECT [chr]
      ,[snp]
      ,[bp]
      ,[pvalhol]
      ,[pvalusa]
      ,[newpvalueholusa]
      ,[pvalfra]
      ,[newpvalueholfrausa]
      ,[pvalbos]
      ,[newpvaluebosholfrausa]
      ,[pvaluk]
      ,[newpvalueukbosholfrausa]
      ,[newpvalueholfra]
      ,[newpvalueboshol]
      ,[newpvalueukhol]
      ,[newpvaluebosholusa]
      ,[newpvaluebosholfra]
      ,[newpvalueukholusa]
      ,[newpvalueukholfra]
      ,[newpvalueukbosusa]
      ,[newpvalueukboshol]
      ,[newpvalueukbosholfra]
      ,[newpvalueukbosholusa]
      ,[newpvalueukholfrausa]
      ,[id]
      ,[log_holusa]
      ,[log_holfrausa]
      ,[log_bosholfrausa]
      ,[log_ukbosholfrausa]
      ,[log_holfra]
      ,[log_boshol]
      ,[log_ukhol]
      ,[log_bosholusa]
      ,[log_bosholfra]
      ,[log_ukholusa]
      ,[log_ukholfra]
      ,[log_ukbosusa]
      ,[log_ukboshol]
      ,[log_ukbosholfra]
      ,[log_ukbosholusa]
      ,[log_ukholfrausa]
  FROM [Alsod].[dbo].[GWA_UK_BOS_HOL_FRA_NIH]
```

GWA_UK_BOS_HOL_FRA_USA

```
SELECT [chr]
      ,[snp]
      ,[bp]
```

```
      ,[pvalhol]
      ,[pvalusa]
      ,[newpvalueholusa]
      ,[pvalfra]
      ,[newpvalueholfrausa]
      ,[pvalbos]
      ,[newpvaluebosholfrausa]
      ,[pvaluk]
      ,[newpvalueukbosholfrausa]
      ,[newpvalueholfra]
      ,[newpvalueboshol]
      ,[newpvalueukhol]
      ,[newpvaluebosholusa]
      ,[newpvaluebosholfra]
      ,[newpvalueukholusa]
      ,[newpvalueukholfra]
      ,[newpvalueukbosusa]
      ,[newpvalueukboshol]
      ,[newpvalueukbosholfra]
      ,[newpvalueukbosholusa]
      ,[newpvalueukholfrausa]
      ,[id]
      ,[log_holusa]
      ,[log_holfrausa]
      ,[log_bosholfrausa]
      ,[log_ukbosholfrausa]
      ,[log_holfra]
      ,[log_boshol]
      ,[log_ukhol]
      ,[log_bosholusa]
      ,[log_bosholfra]
      ,[log_ukholusa]
      ,[log_ukholfra]
      ,[log_ukbosusa]
      ,[log_ukboshol]
      ,[log_ukbosholfra]
      ,[log_ukbosholusa]
      ,[log_ukholfrausa]
      ,[neg2logpvalhol]
      ,[neg2logpvalusa]
      ,[neg2logpvalfra]
      ,[neg2logpvalbos]
      ,[neg2logpvaluk]
  FROM [Alsod].[dbo].[GWA_UK_BOS_HOL_FRA_USA]
```

GWA_CATALOGUE

```
SELECT [Added_date]
      ,[Pubmed_id]
      ,[First_author]
      ,[Submission_date]
      ,[Journal]
      ,[Link]
      ,[Study]
      ,[Disease_trait]
      ,[Initial_sample_size]
      ,[Replication_sample_size]
      ,[Region]
      ,[Reported_genes]
      ,[Strongest_snps_risk_allele]
      ,[Snps]
      ,[Risk_allele_frequency]
      ,[P_value]
```

```
      ,[P_value_text]
      ,[Odd_ratio_beta]
      ,[CI_95_percent_text]
      ,[Platform_snp_passing_QC]
      ,[CNV]
      ,[S_No]
  FROM [Alsod].[dbo].[GWA_CATALOGUE]
```

GWA_ALL_COMMON_DATA

```
SELECT [id]
      ,[chr]
      ,[snp]
      ,[bp]
      ,[pvalbos]
      ,[pvalhol]
      ,[pvaluk]
      ,[new_pvalue_bos]
      ,[new_pvalue_hol]
      ,[new_pvalue_uk]
      ,[new_pvalue_sum]
  FROM [Alsod].[KCLAD\spngoka].[GWA_ALL_COMMON_DATA]
```

GWA_CHIDIST_ALL_COMMON

```
SELECT [id]
      ,[newpvalue]
      ,[chidist]
      ,[pvalue_plot]
  FROM [Alsod].[KCLAD\spngoka].[GWA_CHIDIST_ALL_COMMON]
```

GWASTUDY

```
SELECT [ID]
      ,[POP]
      ,[POPCODE]
      ,[STUDY]
      ,[YEAR]
      ,[FROMNUM]
      ,[TONUM]
      ,[LINKID]
  FROM [Alsod].[KCLAD\spngoka].[GWASTUDY]
```

GWA_FOGH

```
SELECT [SNP]
      ,[CHROMOSOME]
      ,[BP]
      ,[PVALUE]
  FROM [Alsod].[dbo].[GWA_FOGH]
```

Institution

```
SELECT [institution_code]
      ,[institution_name]
      ,[department]
      ,[institution_address1]
      ,[institution_address2]
      ,[institution_address3]
      ,[city]
      ,[state]
      ,[postcode]
      ,[country_iso_code]
```

```
      ,[contact_title]
      ,[contact_first_name]
      ,[contact_middle_init]
      ,[contact_last_name]
      ,[telephone]
      ,[extension]
      ,[facsimile]
      ,[e_mail_address]
  FROM [Alsod].[dbo].[institution]
```

IPAddress of countries

```
SELECT [ip1]
      ,[ip2]
      ,[registry]
      ,[assigned]
      ,[iso2]
      ,[iso3]
      ,[country]
  FROM [Alsod].[dbo].[ipcountry]
```

Members

```
SELECT [member_id]
      ,[contact_title]
      ,[contact_first_name]
      ,[contact_middle_init]
      ,[contact_last_name]
      ,[institution_name]
      ,[department]
      ,[institution_address1]
      ,[institution_address2]
      ,[institution_address3]
      ,[city]
      ,[postcode]
      ,[country_iso_code]
      ,[telephone]
      ,[extension]
      ,[facsimile]
      ,[e_mail_address]
  FROM [Alsod].[dbo].[members]
```

NETPHOS_scores

```
SELECT [gene]
      ,[codon]
      ,[sequence]
      ,[score]
      ,[phosphorylated]
      ,[message]
  FROM [Alsod].[dbo].[NETPHOS_scores]
```

No_PATHOGENICITY

```
SELECT [Gene]
      ,[Pathogenic]
  FROM [Alsod].[dbo].[No_PATHOGENICITY]
```

No_SALS

```
SELECT [Gene]
      ,[Total_no_family_history]
  FROM [Alsod].[dbo].[No_SALS]
```

PANTHER_scores

```
SELECT [subPSEC]
```

```
      ,[P_deletrious]
      ,[gene]
      ,[mutation]
      ,[codon]
      ,[name]
      ,[Accession]
      ,[score]
      ,[position]
      ,[Pwt]
      ,[Psubstituted]
      ,[NIC]
      ,[message]
  FROM [Alsod].[dbo].[PANTHER_scores]
```

POLYPHEN_scores

```
SELECT [gene]
      ,[mutation]
      ,[codon]
      ,[position]
      ,[PSIC_score]
      ,[prediction]
      ,[PSIC_score_new]
  FROM [Alsod].[dbo].[POLYPHEN_scores]
```

SIFT_scores

```
SELECT [position]
      ,[codon]
      ,[gene]
      ,[mutation_from]
      ,[mutation_to]
      ,[message]
      ,[subPSEC]
  FROM [Alsod].[dbo].[SIFT_scores]
```

Population_frequency_1000genome

```
SELECT [mutation_id]
      ,[gene_id]
      ,[sequence_original_mutated]
      ,[sequence_original]
      ,[sequence_mutated]
      ,[Location]
      ,[Ref_Allele]
      ,[Var_Allele]
      ,[Global_Count]
      ,[Global_Freq]
      ,[South_American_Count]
      ,[South_American_Freq]
      ,[Asian_Count]
      ,[Asian_Freq]
      ,[African_Count]
      ,[African_Freq]
      ,[European_Count]
      ,[European_Freq]
  FROM [Alsod].[dbo].[population_frequency_1000genome]
```

Population_frequency_EVS

```
SELECT [mutation_id]
      ,[gene_id]
      ,[sequence_original_mutated]
      ,[sequence_original]
      ,[sequence_mutated]
      ,[Location]
```

```
      ,[Ref_Allele]
      ,[Var_Allele]
      ,[Global_Count]
      ,[Global_Freq]
      ,[European_American_Count]
      ,[European_American_Freq]
      ,[African_American_Count]
      ,[African_American_Freq]
  FROM [Alsod].[dbo].[population_frequency_EVS]
```

Random_subpsec scores

```
SELECT [gene_id]
      ,[mutation_mnemonic]
      ,[mutated_from_to]
      ,[exon_number]
      ,[codon_id]
      ,[subpsec_score]
  FROM [Alsod].[dbo].[random_subpsec]
```

Reference_MGI

```
SELECT [MGI_id]
      ,[gene_id]
      ,[pubmed_id]
      ,[ID]
  FROM [Alsod].[dbo].[Reference_MGI]
```

Reference_Pubmed

```
SELECT [title]
      ,[pubmed_id]
      ,[year]
      ,[first_author]
      ,[ID]
  FROM [Alsod].[dbo].[Reference_Pubmed]
```

Refs

```
SELECT [reference_id]
      ,[reference_pubmed_id]
      ,[reference_medline_id]
      ,[reference]
      ,[reference_title]
      ,[reference_authors]
      ,[reference_year]
      ,[reference_type]
      ,[reference_subtype]
      ,[reference_comment1]
      ,[reference_comment2]
      ,[reference_comment3]
  FROM [Alsod].[dbo].[refs]
```

Refs_lookup

```
SELECT [reference_id]
      ,[mutation_id]
      ,[gene_id]
      ,[patient_id]
  FROM [Alsod].[dbo].[refs_lookup]
```

Survey

```
SELECT [id]
      ,[age]
      ,[gender]
      ,[country]
      ,[status]
```

```
      ,[q211]
      ,[q212]
      ,[q220]
      ,[q230]
      ,[q240]
      ,[q250]
      ,[q310]
      ,[q320]
      ,[q330]
      ,[q340]
      ,[q350]
      ,[q360]
      ,[q371]
      ,[q372]
      ,[q373]
      ,[q374]
      ,[q375]
      ,[q410]
      ,[q420]
      ,[q430]
      ,[q510]
      ,[q520]
      ,[q530]
      ,[q540]
      ,[q550]
      ,[referrer]
      ,[other_referrer]
  FROM [Alsod].[dbo].[survey]
```

Trinucleotides

```
SELECT [trinucleotide]
      ,[amino_acid]
      ,[code]
  FROM [Alsod].[dbo].[trinucleotides]
```

UNIPROT_Mutation

```
SELECT [SN]
      ,[ORIGINAL]
      ,[POSITION]
      ,[VARIATION]
      ,[MUTATION]
      ,[ID]
      ,[TYPE]
      ,[DESCRIPTION]
      ,[DATA_FROM]
  FROM [Alsod].[dbo].[UNIPROT_MUTATION]
```

Users

```
SELECT [user_id]
      ,[institution_code]
      ,[username]
      ,[password]
      ,[superuser]
      ,[user_email]
      ,[contact_title]
      ,[contact_first_name]
      ,[contact_middle_init]
      ,[contact_last_name]
      ,[institution_name]
      ,[department]
      ,[institution_address1]
      ,[institution_address2]
```

```
      ,[institution_address3]
      ,[state]
      ,[city]
      ,[postcode]
      ,[country_iso_code]
      ,[telephone]
      ,[facsimile]
      ,[contact_person]
      ,[extension]
  FROM [Alsod].[dbo].[users]
```

Version

```
SELECT [major]
      ,[minor]
      ,[langid]
      ,[sublangid]
      ,[sortid]
  FROM [Alsod].[dbo].[version]
```

Visitors

```
SELECT [DateTime]
      ,[IPAddress]
      ,[Host]
      ,[Country]
      ,[Page]
      ,[summaryhttp]
      ,[summaryraw]
      ,[ID]
      ,[CountryCode]
  FROM [Alsod].[dbo].[visitors]
```

genome

```
SELECT [structure_id]
      ,[chrom]
      ,[strand]
      ,[txtStart]
      ,[txtEnd]
      ,[cdsStart]
      ,[cdsEnd]
      ,[exonCount]
      ,[exonStarts]
      ,[exonEnds]
      ,[proteinID]
      ,[alignID]
      ,[id]
  FROM [Alsod].[KCLAD\spngoka].[genome]
```

Appendix 25 – T-SQL scripts in each SQL view

ANALYZE_JOINTDATA_VIEW

```
SELECT     dbo.ANALYZE_THEIRDATA.CHROMOSOME, dbo.ANALYZE_THEIRDATA.SNP,
dbo.ANALYZE_THEIRDATA.BP, dbo.ANALYZE_THEIRDATA.PVALUE,
                      dbo.ANALYZE_OURDATA.PVALHOL,
dbo.ANALYZE_OURDATA.PVALUSA, dbo.ANALYZE_OURDATA.PVALFRA,
dbo.ANALYZE_OURDATA.PVALBOS,
                      dbo.ANALYZE_OURDATA.PVALUK,
dbo.ANALYZE_THEIRDATA.neg2logpval, dbo.ANALYZE_OURDATA.neg2logpvalhol,
                      dbo.ANALYZE_OURDATA.neg2logpvalusa,
dbo.ANALYZE_OURDATA.neg2logpvalfra, dbo.ANALYZE_OURDATA.neg2logpvalbos,
                      dbo.ANALYZE_OURDATA.neg2logpvaluk
FROM         dbo.ANALYZE_THEIRDATA INNER JOIN
                      dbo.ANALYZE_OURDATA ON dbo.ANALYZE_THEIRDATA.SNP =
dbo.ANALYZE_OURDATA.SNP
```

countrydetails_view

```
SELECT     dbo.countryinfo.country_iso_code, dbo.countryinfo.country,
dbo.countryinfo.fips104, dbo.countryinfo.iso2, dbo.countryinfo.iso3,
dbo.countryinfo.isono,
                      dbo.countryinfo.capital, dbo.countryinfo.region,
dbo.countryinfo.currency, dbo.countryinfo.currency_code,
dbo.countryinfo.population,
                      dbo.countryinfo.t_lag, dbo.countryinfo.t_type,
dbo.countryinfo.comments, dbo.countryinfo.language,
dbo.countryinfo.city_info,
                      dbo.countryinfo.dialing_code,
dbo.countrylatlong.latitude, dbo.countrylatlong.longitude
FROM         dbo.countryinfo INNER JOIN
                      dbo.countrylatlong ON
dbo.countryinfo.country_iso_code = dbo.countrylatlong.country_iso_code
```

gender_report

```
SELECT     dbo.patient_genetic_record.gene_id AS Gene, COUNT(CASE [sex]
WHEN 'female' THEN 'Others' ELSE NULL END) AS Male,
                      COUNT(CASE [sex] WHEN 'male' THEN 'Others' ELSE NULL
END) AS Female, COUNT(CASE [sex] WHEN 'Others' THEN 'Others' ELSE NULL END)
                      AS Others, COUNT(dbo.patient_record.sex) AS Total
FROM         dbo.patient_genetic_record INNER JOIN
                      dbo.patient_record ON
dbo.patient_genetic_record.patient_id = dbo.patient_record.patient_id
GROUP BY dbo.patient_genetic_record.gene_id
```

mutationtype_report

```
SELECT     dbo.patient_genetic_record.gene_id AS Gene, COUNT(CASE
[mutation_type] WHEN 'Compound' THEN 'Others' ELSE NULL END) AS Compound,
                      COUNT(CASE [mutation_type] WHEN 'Substitution' THEN
'Others' ELSE NULL END) AS Substitution,
                      COUNT(CASE [mutation_type] WHEN 'Deletion' THEN
'Others' ELSE NULL END) AS Deletion,
                      COUNT(CASE [mutation_type] WHEN 'Polymorphism' THEN
'Others' ELSE NULL END) AS Polymorphism,
                      COUNT(CASE [mutation_type] WHEN 'Insertion' THEN
'Others' ELSE NULL END) AS Insertion,
```

```
                      COUNT(CASE [mutation_type] WHEN 'Others' THEN
'Others' ELSE NULL END) AS Others, COUNT(dbo.mutation.mutation_type) AS
Total
FROM         dbo.patient_genetic_record INNER JOIN
                      dbo.mutation ON
dbo.patient_genetic_record.mutation_id = dbo.mutation.mutation_id AND
                      dbo.patient_genetic_record.mutation_id =
dbo.mutation.mutation_id
```

patientmutationrecords

```
SELECT     COUNT(dbo.patient_genetic_record.gene_id) AS SODCOMPOUND
FROM         dbo.patient_record INNER JOIN
                      dbo.patient_genetic_record ON
dbo.patient_record.patient_id = dbo.patient_genetic_record.patient_id INNER
JOIN
                      dbo.mutation ON
dbo.patient_genetic_record.mutation_id = dbo.mutation.mutation_id AND
                      dbo.patient_genetic_record.mutation_id =
dbo.mutation.mutation_id
WHERE     (dbo.patient_genetic_record.gene_id = 'sod1') AND
(dbo.mutation.mutation_type = 'compound')
```

site_of_onset_report

```
SELECT     dbo.patient_genetic_record.gene_id AS Gene, COUNT(CASE
[site_of_onset] WHEN 'leg' THEN 'Others' ELSE NULL END) AS Leg,
                      COUNT(CASE [site_of_onset] WHEN 'arm' THEN 'Others'
ELSE NULL END) AS Arm,
                      COUNT(CASE [site_of_onset] WHEN 'bulbar' THEN
'Others' ELSE NULL END) AS Bulbar,
                      COUNT(CASE [site_of_onset] WHEN 'Others' THEN
'Others' ELSE NULL END) AS Others, COUNT(dbo.patient_record.site_of_onset)
AS Total
FROM         dbo.patient_genetic_record INNER JOIN
                      dbo.patient_record ON
dbo.patient_genetic_record.patient_id = dbo.patient_record.patient_id
GROUP BY dbo.patient_genetic_record.gene_id
```

SOD1_MUTATIONS

```
SELECT     TOP (100) PERCENT mutation_mnemonic, gene_id, mutation_type,
codon, sequence_location_number
FROM         dbo.mutation
WHERE     (gene_id = 'SOD1') AND (codon <> 0)
```

statistics_country

```
SELECT     TOP (100) PERCENT Country, COUNT(*) AS Frequency
FROM         dbo.visitors
GROUP BY Country
ORDER BY Frequency
```

study_expression

```
SELECT     TOP (100) PERCENT dbo.gene_study.Lastname + ',' +
dbo.gene_study.Year + ' [' + dbo.gene_study.gene_id + ']' AS expression,
```

```
                      dbo.gene_study.Lastname, dbo.gene_study.Month,
dbo.gene_study.Year, dbo.gene_study.Title, dbo.gene_study.PMID,
dbo.gene.gene_id,
                      dbo.gene_study.StudyID, dbo.gene.ucsc_id,
dbo.gene.protein_name, dbo.gene.other_names,
dbo.gene.reason_for_investigation, dbo.gene.result,
                      dbo.gene_study.Others, dbo.gene_study.Volume,
dbo.gene.ncbi_locuslink_id, dbo.gene.structure_id, dbo.gene.omim_id,
dbo.gene.chromosome,
                      dbo.gene.category, dbo.gene.hgnc_id,
dbo.gene_study.Comment, dbo.gene_study.doi_key, dbo.gene_study.Paperlink,
dbo.gene.keywords,
                      dbo.gene.swissport_id, dbo.gene.ihop,
dbo.gene.gene_effect
FROM         dbo.gene_study INNER JOIN
                      dbo.gene ON dbo.gene_study.gene_id = dbo.gene.gene_id
```

AllGenes_Trinuc_Codes

```
SELECT     dbo.codons.gene_id, dbo.codons.codon_id AS codon,
dbo.codons.trinucleotide, dbo.trinucleotides.code
FROM         dbo.codons INNER JOIN
                      dbo.trinucleotides ON dbo.codons.trinucleotide =
dbo.trinucleotides.trinucleotide
```

AllMutation_Trinuc_Codes

```
SELECT DISTINCT
                      TOP (100) PERCENT dbo.mutation.gene_id,
dbo.mutation.aa_mutated, dbo.mutation.aa_original, dbo.mutation.codon,
dbo.trinucleotides.code,
                      dbo.mutation.mutation_type
FROM         dbo.mutation INNER JOIN
                      dbo.trinucleotides ON dbo.mutation.aa_mutated =
dbo.trinucleotides.amino_acid
ORDER BY dbo.mutation.codon
```

Credibility_score_analysis

```
SELECT     [KCLAD\spngoka].Number_of_mutations_per_gene.Rank AS
Rank_Mutations,

[KCLAD\spngoka].Number_of_affected_patients_in_ALSoD.Rank AS Rank_Patients,
                      [KCLAD\spngoka].Number_of_cases_affected.Rank AS
Rank_Affected_Cases,
                      [KCLAD\spngoka].Number_of_cases_tested.Rank AS
Rank_Tested_Cases,
                      [KCLAD\spngoka].Number_of_controls_affected.Rank AS
Rank_Affected_Controls,
                      [KCLAD\spngoka].Number_of_controls_tested.Rank AS
Rank_Tested_Controls,
                      [KCLAD\spngoka].[Number_of_mutations_in
_same_codon_by_rank].Rank AS Rank_Codon,

[KCLAD\spngoka].Number_of_patients_with_family_history_FALS.Rank AS
Rank_FALS,

[KCLAD\spngoka].Number_of_patients_without_family_history_SALS.Rank AS
Rank_SALS,
```

```
[KCLAD\spngoka].Number_of_times_mutation_is_replicated.Rank AS
Rank_Replications,

[KCLAD\spngoka].Number_of_unique_countries_on_genes.Rank AS
Rank_Populations,
                      [KCLAD\spngoka].Ranking_LOD_score.Rank AS
Rank_LOD_Score, [KCLAD\spngoka].Number_of_mutations_per_gene.Gene
FROM         [KCLAD\spngoka].Number_of_affected_patients_in_ALSoD INNER
JOIN
                      [KCLAD\spngoka].Number_of_mutations_per_gene ON

[KCLAD\spngoka].Number_of_affected_patients_in_ALSoD.Gene =
[KCLAD\spngoka].Number_of_mutations_per_gene.Gene INNER JOIN
                      [KCLAD\spngoka].Number_of_cases_affected ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Number_of_cases_affected.Gene INNER JOIN
                      [KCLAD\spngoka].Number_of_cases_tested ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Number_of_cases_tested.Gene INNER JOIN
                      [KCLAD\spngoka].Number_of_controls_affected ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Number_of_controls_affected.Gene INNER JOIN
                      [KCLAD\spngoka].Number_of_controls_tested ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Number_of_controls_tested.Gene INNER JOIN
                      [KCLAD\spngoka].[Number_of_mutations_in
_same_codon_by_rank] ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].[Number_of_mutations_in _same_codon_by_rank].Gene INNER
JOIN

[KCLAD\spngoka].Number_of_patients_with_family_history_FALS ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Number_of_patients_with_family_history_FALS.Gene INNER JOIN

[KCLAD\spngoka].Number_of_patients_without_family_history_SALS ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Number_of_patients_without_family_history_SALS.Gene INNER
JOIN

[KCLAD\spngoka].Number_of_times_mutation_is_replicated ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Number_of_times_mutation_is_replicated.Gene INNER JOIN
                      [KCLAD\spngoka].Number_of_unique_countries_on_genes
ON
                      [KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Number_of_unique_countries_on_genes.Gene INNER JOIN
                      [KCLAD\spngoka].Ranking_LOD_score ON
[KCLAD\spngoka].Number_of_mutations_per_gene.Gene =
[KCLAD\spngoka].Ranking_LOD_score.gene
```

daily_visitors_report

```
SELECT     TOP (100) PERCENT LEFT(DateTime, 12) AS Daily, ID
FROM         dbo.visitors
ORDER BY ID
```

Gridview_Replicated_mutations

```
SELECT     Gene, mutation_mnemonic AS Mutation, COUNT(mutation_mnemonic) AS
Frequency, DENSE_RANK() OVER (ORDER BY COUNT(mutation_mnemonic) DESC)
 AS Rank_Mutation
FROM          (SELECT     [KCLAD\spngoka].patient_record_view.Gene,
dbo.mutation.mutation_mnemonic,
[KCLAD\spngoka].patient_record_view.first_author,

[KCLAD\spngoka].patient_record_view.year,
[KCLAD\spngoka].patient_record_view.pubmed_id
                      FROM          [KCLAD\spngoka].patient_record_view
INNER JOIN
                                                dbo.mutation ON
[KCLAD\spngoka].patient_record_view.mid = dbo.mutation.mutation_id
                      UNION
                      SELECT DISTINCT gene_id, mutation_mnemonic,
First_Author, Year, pubmed_id
                      FROM          dbo.gene_SNP_lifeseq
                      WHERE     (mutation_mnemonic IS NOT NULL) AND
(mutation_mnemonic <> 'NA')
                      UNION
                      SELECT DISTINCT gene, mutation, author, year,
pubmed_id
                      FROM          dbo.mutation_replication
                      UNION
                      SELECT DISTINCT gene_id, mutation_mnemonic,
First_Author, Year, pubmed_id
                      FROM          dbo.gene_SNP_ALSGene
                      WHERE     (mutation_mnemonic IS NOT NULL) AND
(mutation_mnemonic <> 'NA')) AS derivedtbl_1
```

Gridview_total_cases

```
SELECT DISTINCT TOP (100) PERCENT gene, cases, author, year, pubmed_id
FROM          dbo.mutation_replication
WHERE     (cases IS NOT NULL)
ORDER BY gene
```

Gridview_total_controls

```
SELECT DISTINCT gene, cases, controls, author, year, pubmed_id
FROM          dbo.mutation_replication
WHERE     (controls IS NOT NULL)
```

Gridview_unique_countries

```
SELECT DISTINCT gene_id, country
FROM          dbo.gene_SNP_lifeseq
WHERE     (family_history IS NOT NULL) AND (gender IS NOT NULL) AND
(mutation_mnemonic IS NOT NULL) AND (country IS NOT NULL)
UNION
SELECT DISTINCT Gene, Country
FROM          [KCLAD\spngoka].patient_record_view
GROUP BY Gene, Country
```

GWA_ALL

```
SELECT     [KCLAD\spngoka].GWA_BOS.S_No AS snbos,
[KCLAD\spngoka].GWA_HOL.S_No AS snhol, [KCLAD\spngoka].GWA_UK.S_No AS snuk,
```

```
                      [KCLAD\spngoka].GWA_BOS.CHROMOSOME AS chr,
[KCLAD\spngoka].GWA_BOS.SNP AS snpbos, [KCLAD\spngoka].GWA_HOL.SNP AS
snphol,
                      [KCLAD\spngoka].GWA_UK.SNP AS snpuk,
[KCLAD\spngoka].GWA_BOS.BP AS bpbos, [KCLAD\spngoka].GWA_HOL.BP AS bphol,
                      [KCLAD\spngoka].GWA_UK.BP AS bpuk,
[KCLAD\spngoka].GWA_BOS.PVALUE AS pvalbos, [KCLAD\spngoka].GWA_HOL.PVALUE
AS pvalhol,
                      [KCLAD\spngoka].GWA_UK.PVALUE AS pvaluk
FROM         [KCLAD\spngoka].GWA_BOS INNER JOIN
                      [KCLAD\spngoka].GWA_HOL ON
[KCLAD\spngoka].GWA_BOS.SNP = [KCLAD\spngoka].GWA_HOL.SNP INNER JOIN
                      [KCLAD\spngoka].GWA_UK ON [KCLAD\spngoka].GWA_BOS.SNP
= [KCLAD\spngoka].GWA_UK.SNP
```

GWA_ALL_CHIDIST_PLOT

```
SELECT     [KCLAD\spngoka].GWA_ALL_COMMON_DATA.id,
[KCLAD\spngoka].GWA_CHIDIST_ALL_COMMON.id AS id2,
                      [KCLAD\spngoka].GWA_ALL_COMMON_DATA.chr,
[KCLAD\spngoka].GWA_ALL_COMMON_DATA.snp,
[KCLAD\spngoka].GWA_ALL_COMMON_DATA.bp,
                      [KCLAD\spngoka].GWA_ALL_COMMON_DATA.pvalbos,
[KCLAD\spngoka].GWA_ALL_COMMON_DATA.pvalhol,
                      [KCLAD\spngoka].GWA_ALL_COMMON_DATA.pvaluk,
[KCLAD\spngoka].GWA_CHIDIST_ALL_COMMON.newpvalue,
                      [KCLAD\spngoka].GWA_CHIDIST_ALL_COMMON.chidist,
[KCLAD\spngoka].GWA_CHIDIST_ALL_COMMON.pvalue_plot
FROM         [KCLAD\spngoka].GWA_ALL_COMMON_DATA INNER JOIN
                      [KCLAD\spngoka].GWA_CHIDIST_ALL_COMMON ON
                      [KCLAD\spngoka].GWA_ALL_COMMON_DATA.id =
[KCLAD\spngoka].GWA_CHIDIST_ALL_COMMON.id
```

latitude_longitude_mutation

```
SELECT     TOP (100) PERCENT dbo.mutation.mutation_mnemonic,
dbo.countrydetails.latitude, dbo.countrydetails.longitude,
dbo.countrydetails.country
FROM         dbo.mutation INNER JOIN
                      dbo.countrydetails ON dbo.mutation.countryiso1 =
dbo.countrydetails.country_iso_code OR
                      dbo.mutation.countryiso2 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso3 =
dbo.countrydetails.country_iso_code OR
                      dbo.mutation.countryiso4 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso5 =
dbo.countrydetails.country_iso_code OR
                      dbo.mutation.countryiso6 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso7 =
dbo.countrydetails.country_iso_code OR
                      dbo.mutation.countryiso8 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso9 =
dbo.countrydetails.country_iso_code OR
                      dbo.mutation.countryiso10 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso11 =
dbo.countrydetails.country_iso_code OR
                      dbo.mutation.countryiso12 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso13 =
dbo.countrydetails.country_iso_code OR
```

```
                        dbo.mutation.countryiso14 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso15 =
dbo.countrydetails.country_iso_code OR
                        dbo.mutation.countryiso16 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso17 =
dbo.countrydetails.country_iso_code OR
                        dbo.mutation.countryiso18 =
dbo.countrydetails.country_iso_code OR dbo.mutation.countryiso19 =
dbo.countrydetails.country_iso_code OR
                        dbo.mutation.countryiso20 =
dbo.countrydetails.country_iso_code
WHERE     (dbo.mutation.gene_id = 'SOD1') AND
(dbo.mutation.sequence_location_type = 'exon')
ORDER BY dbo.mutation.codon
```

mutation update

```
SELECT     mutation_alias, mutation_name, mutation_mnemonic, gene_id,
mutation_type, codon, first_author, year, countryiso1, pubmed_id, doi
FROM         dbo.mutation
WHERE     (gene_id = 'ALS2')
```

No_SALS_view

```
SELECT     Gene, COUNT([family History]) AS Total_no_family_history
FROM         (SELECT     [KCLAD\spngoka].patient_record_view.sex,
[KCLAD\spngoka].patient_record_view.Country,

[KCLAD\spngoka].patient_record_view.[family History],
[KCLAD\spngoka].patient_record_view.Gene, dbo.mutation.mutation_mnemonic,

[KCLAD\spngoka].patient_record_view.first_author,
[KCLAD\spngoka].patient_record_view.year,

[KCLAD\spngoka].patient_record_view.pubmed_id,
[KCLAD\spngoka].patient_record_view.[Site of Onset]
                        FROM          [KCLAD\spngoka].patient_record_view
INNER JOIN
                                                dbo.mutation ON
[KCLAD\spngoka].patient_record_view.Mutation = dbo.mutation.mutation_name
                        WHERE
([KCLAD\spngoka].patient_record_view.[family History] = 'No')
                        UNION
                        SELECT     gender, country, family_history, gene_id,
mutation_mnemonic, First_Author, Year, pubmed_id, Site_onset
                        FROM         dbo.gene_SNP_lifeseq
                        WHERE     (family_history IS NOT NULL) AND (gender
IS NOT NULL) AND (mutation_mnemonic IS NOT NULL) AND (family_history =
'No'))
                      AS table1
GROUP BY Gene
```

Number_of_affected_patients_in_ALSoD

```
SELECT     gene_id AS Gene, COUNT(*) AS Frequency, DENSE_RANK() OVER (ORDER
BY COUNT(gene_id) DESC) AS Rank
FROM         dbo.patient_genetic_record
WHERE     (gene_id <> 'LUM') AND (gene_id <> 'CRYM')
GROUP BY gene_id
```

Number_of_cases_affected

```
SELECT     gene_id AS Gene, SUM(rate_patients_mutation_found) AS
Total_cases_affected, DENSE_RANK() OVER (ORDER BY
SUM(rate_patients_mutation_found)
                      DESC) AS Rank
FROM         (SELECT DISTINCT gene_id, rate_patients_mutation_found,
First_Author, Year
                      FROM          dbo.gene_SNP_lifeseq
                      WHERE      (rate_patients_mutation_found IS NOT
NULL)) AS table1
GROUP BY gene_id
```

Number_of_cases_recorded

```
SELECT     gene AS Gene, SUM(cases) AS Total_cases_recorded, DENSE_RANK()
OVER (ORDER BY SUM(cases) DESC) AS Rank
FROM         (SELECT DISTINCT gene, cases, controls, author, year,
pubmed_id
                      FROM          dbo.mutation_replication
                      WHERE      (cases IS NOT NULL)) AS table1
GROUP BY gene
```

Number_of_cases_tested

```
SELECT     gene_id AS Gene, SUM(rate_patients_mutation_not_found) AS
Total_cases_tested, DENSE_RANK() OVER (ORDER BY
SUM(rate_patients_mutation_not_found)
                      DESC) AS Rank
FROM         (SELECT DISTINCT gene_id, First_Author, Year,
rate_patients_mutation_not_found
                      FROM          dbo.gene_SNP_lifeseq
                      WHERE      (rate_patients_mutation_not_found IS NOT
NULL)) AS table1
GROUP BY gene_id
```

Number_of_controls_affected

```
SELECT     gene_id AS Gene, SUM(rate_controls_mutation_found) AS
Total_controls_affected, DENSE_RANK() OVER (ORDER BY
SUM(rate_controls_mutation_found)
                      ASC) AS Rank
FROM         (SELECT DISTINCT gene_id, First_Author, Year,
rate_controls_mutation_found
                      FROM          dbo.gene_SNP_lifeseq
                      WHERE      (rate_controls_mutation_found IS NOT
NULL)) AS table1
GROUP BY gene_id
```

Number_of_controls_recorded

```
SELECT     gene AS Gene, SUM(controls) AS Total_controls_recorded,
DENSE_RANK() OVER (ORDER BY SUM(controls) DESC) AS Rank
FROM         (SELECT DISTINCT gene, cases, controls, author, year,
pubmed_id
                      FROM          dbo.mutation_replication
                      WHERE      (controls IS NOT NULL)) AS table1
GROUP BY gene
```

Number_of_controls_tested

```
SELECT     gene_id AS Gene, SUM([rate_controls_mutation_not found]) AS
Total_controls_tested, DENSE_RANK() OVER (ORDER BY
SUM([rate_controls_mutation_not found])
                      DESC) AS Rank
FROM         (SELECT DISTINCT gene_id, [rate_controls_mutation_not found]
                      FROM          dbo.gene_SNP_lifeseq
                      WHERE      ([rate_controls_mutation_not found] IS
NOT NULL)) AS table1
GROUP BY gene_id
```

Number_of_families_with_generations

```
SELECT DISTINCT gene, mutation, codon, score, families, generations,
author, year, pubmed_id, DENSE_RANK() OVER (ORDER BY families DESC)
AS Rank
FROM         dbo.mutation_replication
WHERE     (families IS NOT NULL) AND (generations IS NOT NULL) AND
(families <> '0')
```

Number_of_mutations_in _same_codon_by_rank

```
SELECT     Gene, Sum(Rank) AS Rank_Summation, DENSE_RANK() OVER (ORDER BY
Sum(Rank) DESC) AS Rank
FROM         (SELECT     TOP (100) PERCENT gene_id AS Gene, codon AS Codon,
COUNT(codon) AS Frequency, DENSE_RANK() OVER (ORDER BY COUNT(codon) DESC)
                      AS Rank
FROM         (SELECT     TOP (100) PERCENT gene_id, mutation_mnemonic,
codon
                      FROM          dbo.mutation
                      WHERE      (codon IS NOT NULL) AND (codon <> 0)
                      ORDER BY codon) AS table1
GROUP BY gene_id, codon
ORDER BY codon) AS table2
GROUP BY Gene
```

Number_of_mutations_per_gene

```
SELECT     gene_id AS Gene, COUNT(gene_id) AS Frequency, DENSE_RANK() OVER
(ORDER BY COUNT(gene_id) DESC) AS Rank
FROM         (SELECT     gene_id, mutation_mnemonic
                      FROM          dbo.mutation
                      WHERE      (mutation_mnemonic IS NOT NULL AND
mutation_mnemonic <> 'NA')) AS derivedtbl_1
GROUP BY gene_id
```

Number_of_pathogenic_mutations_by_rank

```
SELECT     Gene, Pathogenic, DENSE_RANK() OVER (ORDER BY ([Pathogenic])
DESC) AS Rank
FROM         (SELECT     Gene, Pathogenic
                      FROM          dbo.No_PATHOGENICITY
                      UNION
                      SELECT     Gene, Pathogenic
                      FROM         [KCLAD\spngoka].PATHOGENICITY_VIEW) AS
table1
```

Number_of_patients_with_family_history_FALS

```
SELECT     Gene, Count([family History]) AS Total_family_history,
DENSE_RANK() OVER (ORDER BY Count([family History]) DESC) AS Rank
FROM         (SELECT     [KCLAD\spngoka].patient_record_view.Sex,
[KCLAD\spngoka].patient_record_view.Country,

[KCLAD\spngoka].patient_record_view.[family History],
[KCLAD\spngoka].patient_record_view.Gene, dbo.mutation.mutation_mnemonic,

[KCLAD\spngoka].patient_record_view.first_author,
[KCLAD\spngoka].patient_record_view.year,

[KCLAD\spngoka].patient_record_view.pubmed_id,
[KCLAD\spngoka].patient_record_view.[Site of Onset]
                      FROM          [KCLAD\spngoka].patient_record_view
INNER JOIN
                                                dbo.mutation ON
[KCLAD\spngoka].patient_record_view.Mutation = dbo.mutation.mutation_name
                      WHERE
([KCLAD\spngoka].patient_record_view.[family History] = 'Yes')
                      UNION
                      SELECT     gender, country, family_history, gene_id,
mutation_mnemonic, First_Author, Year, pubmed_id, Site_onset
                      FROM         dbo.gene_SNP_lifeseq
                      WHERE     (family_history IS NOT NULL) AND (gender
IS NOT NULL) AND (mutation_mnemonic IS NOT NULL) AND (family_history =
'Yes'))
                     AS table1
```

Number_of_patients_without_family_history_SALS

```
SELECT     Gene, Total_no_family_history, DENSE_RANK() OVER (ORDER BY
[Total_no_family_history] DESC) AS Rank
FROM         (SELECT     Gene, Total_no_family_history
                      FROM          dbo.No_SALS
                      UNION
                      SELECT     Gene, Total_no_family_history
                      FROM         [KCLAD\spngoka].No_SALS_view) AS table1
```

Number_of_publications

```
SELECT     gene_id AS Gene, Count(gene_id) AS Total_Publication,
DENSE_RANK() OVER (ORDER BY COUNT(gene_id) DESC) AS Rank
FROM         (SELECT     [Lastname], [Year], [gene_id], [Title], [PMID],
[Paperlink]
                      FROM          dbo.gene_study
                      UNION
                      SELECT DISTINCT first_author, year, gene_id, title,
pubmed_id, link
                      FROM         dbo.mutation
                      WHERE     (first_author <> 'NULL')
                      UNION
                      SELECT DISTINCT First_Author, Year, gene_id,
paperlink, pubmed_id, link
                      FROM         dbo.gene_SNP_updated
                      WHERE     (First_Author <> 'NULL') AND (paperlink <>
'NULL')) AS table1
GROUP BY gene_id
```

Number_of_times_mutation_is_replicated

```
SELECT     Gene, Sum(Rank_Mutation) AS Rank_Summation, DENSE_RANK() OVER
(ORDER BY Sum(Rank_Mutation) DESC) AS Rank
FROM         (SELECT     Gene, mutation_mnemonic AS Mutation,
COUNT(mutation_mnemonic) AS Frequency, DENSE_RANK() OVER (ORDER BY
COUNT(mutation_mnemonic)
                                                  DESC) AS Rank_Mutation
FROM         (SELECT     [KCLAD\spngoka].patient_record_view.Gene,
dbo.mutation.mutation_mnemonic,
[KCLAD\spngoka].patient_record_view.first_author,

[KCLAD\spngoka].patient_record_view.year,
[KCLAD\spngoka].patient_record_view.pubmed_id
                      FROM          [KCLAD\spngoka].patient_record_view
INNER JOIN
                                               dbo.mutation ON
[KCLAD\spngoka].patient_record_view.mid = dbo.mutation.mutation_id
                      UNION
                      SELECT DISTINCT gene_id, mutation_mnemonic,
First_Author, Year, pubmed_id
                      FROM         dbo.gene_SNP_lifeseq
                      WHERE     (mutation_mnemonic IS NOT NULL) AND
(mutation_mnemonic <> 'NA')
                      UNION
                      SELECT DISTINCT gene, mutation, author, year,
pubmed_id
                      FROM         dbo.mutation_replication
                      UNION
                      SELECT DISTINCT gene_id, mutation_mnemonic,
First_Author, Year, pubmed_id
                      FROM         dbo.gene_SNP_ALSGene
                      WHERE     (mutation_mnemonic IS NOT NULL) AND
(mutation_mnemonic <> 'NA')) AS derivedtbl_1
GROUP BY mutation_mnemonic, Gene) AS table2
GROUP BY Gene
```

Number_of_unique_countries_on_genes

```
SELECT     gene_id AS Gene, Count(gene_id) AS Total_Countries, DENSE_RANK()
OVER (ORDER BY COUNT(gene_id) DESC) AS Rank
FROM         (SELECT DISTINCT gene_id, country
                      FROM          dbo.gene_SNP_lifeseq
                      WHERE      (family_history IS NOT NULL) AND (gender
IS NOT NULL) AND (mutation_mnemonic IS NOT NULL) AND (country IS NOT NULL)
                      UNION
                      SELECT DISTINCT Gene, Country
                      FROM         [KCLAD\spngoka].patient_record_view
                      GROUP BY Gene, Country) AS table1
GROUP BY gene_id
```

Pathogenicity_Check

```
SELECT     TOP (100) PERCENT dbo.PANTHER_scores.gene,
dbo.PANTHER_scores.mutation, dbo.PANTHER_scores.codon,
                      dbo.PANTHER_scores.subPSEC AS PANTHER,
dbo.PANTHER_scores.message AS Prediction1,
                      dbo.POLYPHEN_scores.PSIC_score_new AS POLYPHEN,
dbo.POLYPHEN_scores.prediction AS Prediction2, dbo.SIFT_scores.subPSEC AS
SIFT,
```

```
                      dbo.SIFT_scores.message AS Prediction3, CASE WHEN
dbo.PANTHER_scores.message = 'Deletrious' THEN '1' ELSE '0' END AS
Panther_rate,
                      CASE WHEN (dbo.POLYPHEN_scores.prediction = 'Benign'
OR
                      dbo.POLYPHEN_scores.prediction = 'NA') THEN '0' ELSE
'1' END AS Polyphen_rate,
                      CASE WHEN dbo.SIFT_scores.message = 'AFFECT PROTEIN
FUNCTION ' THEN '1' ELSE '0' END AS Sift_rate
FROM         dbo.PANTHER_scores INNER JOIN
                      dbo.POLYPHEN_scores ON dbo.PANTHER_scores.mutation =
dbo.POLYPHEN_scores.mutation AND
                      dbo.PANTHER_scores.gene = dbo.POLYPHEN_scores.gene
INNER JOIN
                      dbo.SIFT_scores ON dbo.POLYPHEN_scores.codon =
dbo.SIFT_scores.codon AND dbo.POLYPHEN_scores.gene = dbo.SIFT_scores.gene
```

Pathogenicity_Check_Summary

```
SELECT     gene, mutation, codon, PANTHER, Prediction1, POLYPHEN,
Prediction2, SIFT, Prediction3, Panther_rate, Polyphen_rate, Sift_rate,
                      CAST(Panther_rate AS int) + CAST(Sift_rate AS int) +
CAST(Polyphen_rate AS int) AS Result, CASE WHEN (CAST(Panther_rate AS int)
                      + CAST(Sift_rate AS int) + CAST(Polyphen_rate AS
int)) >= '1' THEN 'Yes' ELSE 'No' END AS Pathogenic
FROM         [KCLAD\spngoka].Pathogenicity_Check
```

pathogenicity_specific_mutation_stats

```
SELECT     Gene, mutation_mnemonic, COUNT(Gene) AS No_Patients, SUM(CASE
WHEN [family_History] = 'Yes' THEN 1 ELSE 0 END) AS FALS,
                      SUM(CASE WHEN [family_History] = 'No' THEN 1 ELSE 0
END) AS SALS, SUM(CASE WHEN Sex = 'male' THEN 1 ELSE 0 END) AS Male,
                      SUM(CASE WHEN Sex = 'female' THEN 1 ELSE 0 END) AS
Female, SUM(CASE WHEN [Site_of_Onset] = 'leg' THEN 1 ELSE 0 END) AS Leg,
                      SUM(CASE WHEN [Site_of_Onset] = 'arm' THEN 1 ELSE 0
END) AS Arm, SUM(CASE WHEN [Site_of_Onset] = 'limb' THEN 1 ELSE 0 END) AS
Limbs,
                      SUM(CASE WHEN [Site_of_Onset] = 'spinal' THEN 1 ELSE
0 END) AS Spinal, SUM(CASE WHEN [Site_of_Onset] = 'unknown' THEN 1 ELSE 0
END)
                      AS Unknown, SUM(CASE WHEN [Site_of_Onset] = 'bulbar'
THEN 1 ELSE 0 END) AS Bulbar,
                      SUM(CASE WHEN [Site_of_Onset] <> 'bulbar' THEN 1 ELSE
0 END) AS Limb
FROM         (SELECT     Gene, sex, Country, Ethnicity, [Age of Onset] AS
Age_of_Onset, [Site of Onset] AS Site_of_Onset, [family History] AS
family_History,
                                                Mutation, mutation_mnemonic,
first_author, year, pubmed_id, disease_duration
                       FROM          [KCLAD\spngoka].patient_record_view
                       WHERE      (Gene = 'SOD1') AND (mutation_mnemonic =
'I113T')) AS derivedtbl_1
GROUP BY Gene, mutation_mnemonic
```

PATHOGENICITY_VIEW

```
SELECT     gene, COUNT(gene) AS Pathogenic
FROM         (SELECT     gene, mutation, Panther_rate, Sift_rate,
Polyphen_rate, Result, Pathogenic
```

```
                        FROM
[KCLAD\spngoka].Pathogenicity_Check_Summary
                        WHERE       (Pathogenic = 'Yes')) AS derivedtbl_1
GROUP BY gene
```

patient_mutation_country

```
SELECT DISTINCT TOP (100) PERCENT dbo.mutation.mutation_mnemonic,
dbo.Country.country_name, dbo.mutation.codon
FROM         dbo.patient_genetic_record INNER JOIN
                        dbo.patient_record ON
dbo.patient_genetic_record.patient_id = dbo.patient_record.patient_id INNER
JOIN
                        dbo.mutation ON
dbo.patient_genetic_record.mutation_id = dbo.mutation.mutation_id INNER
JOIN
                        dbo.Country ON dbo.patient_record.country_iso_code =
dbo.Country.country_iso_code
WHERE      (dbo.mutation.gene_id = 'SOD1') AND
(dbo.mutation.sequence_location_type = 'exon')
ORDER BY dbo.mutation.codon
```

patient_record_analysis

```
SELECT      Gene, COUNT(mid) AS No_Patients, SUM(CASE WHEN [family History]
= 'Yes' THEN 1 ELSE 0 END) AS FALS,
                        SUM(CASE WHEN [family History] = 'No' THEN 1 ELSE 0
END) AS SALS, SUM(CASE WHEN [family History] = 'Unk' THEN 1 ELSE 0 END) AS
NotKnown,
                        SUM(CASE WHEN Sex = 'male' THEN 1 ELSE 0 END) AS
Male, SUM(CASE WHEN Sex = 'female' THEN 1 ELSE 0 END) AS Female,
                        SUM(CASE WHEN Sex = 'Anonymou' THEN 1 ELSE 0 END) AS
Anonymous, SUM(CASE WHEN [Site of Onset] = 'leg' THEN 1 ELSE 0 END) AS Leg,
                        SUM(CASE WHEN [Site of Onset] = 'arm' THEN 1 ELSE 0
END) AS Arm, SUM(CASE WHEN [Site of Onset] = 'limb' THEN 1 ELSE 0 END) AS
Limb,
                        SUM(CASE WHEN [Site of Onset] = 'spinal' THEN 1 ELSE
0 END) AS Spinal, SUM(CASE WHEN [Site of Onset] = 'unknown' THEN 1 ELSE 0
END)
                        AS Unknown, SUM(CASE WHEN [Site of Onset] = 'bulbar'
THEN 1 ELSE 0 END) AS Bulbar,
                        SUM(CASE WHEN [Site of Onset] <> 'bulbar' THEN 1 ELSE
0 END) AS Non_Bulbar, AVG([Age of Onset]) AS Mean_Age_Onset
FROM         [KCLAD\spngoka].patient_record_view
GROUP BY Gene
```

patient_record_analysis_summary

```
SELECT      SUM(No_Patients) AS Total, SUM(FALS) AS Fals, SUM(SALS) AS Sals,
SUM(NotKnown) AS NotKnown, SUM(Male) AS Male, SUM(Female) AS Female,
                        SUM(Anonymous) AS Anonymous, SUM(Leg) AS Leg,
SUM(Arm) AS Arm, SUM(Limb) AS Limb, SUM(Spinal) AS Spinal, SUM(Unknown) AS
Unknown,
                        SUM(Bulbar) AS Bulbar, SUM(Non_Bulbar) AS Non_Bulbar,
AVG(Mean_Age_Onset) AS [Mean Age]
FROM         [KCLAD\spngoka].patient_record_analysis
```

patient_record_mutation_frequency

```
SELECT     TOP (100) PERCENT Mutation, COUNT(*) AS Frequency, Gene
FROM         [KCLAD\spngoka].patient_record_view
GROUP BY Mutation, Gene
HAVING      (COUNT(*) > 0)
ORDER BY Gene, Frequency DESC
```

patient_record_view

```
SELECT     TOP (100) PERCENT p.sex, c.country_name AS Country,
p.ethnic_origin AS Ethnicity, p.age_of_onset AS [Age of Onset],
                      p.site_of_onset AS [Site of Onset], p.date_of_onset
AS dof, p.alive, g.screened, g.family_history AS [family History],
g.gene_id AS Gene,
                      m.mutation_type, m.mutation_name AS Mutation,
m.codon_id, m.mutation_id AS mid, m.first_author, m.year, m.pubmed_id,
p.disease_duration,
                      m.mutation_mnemonic, g.family_id, CHARINDEX(' ',
g.family_id) AS Space, SUBSTRING(g.family_id, 1, NULLIF (CHARINDEX(' ',
g.family_id) - 1, - 1))
                      AS first_author2, SUBSTRING(g.family_id, CHARINDEX('
', g.family_id) + 1, LEN(g.family_id)) AS year2
FROM         dbo.mutation AS m INNER JOIN
                      dbo.patient_genetic_record AS g ON m.mutation_id =
g.mutation_id INNER JOIN
                      dbo.patient_record AS p ON g.patient_id =
p.patient_id INNER JOIN
                      dbo.Country AS c ON p.country_iso_code =
c.country_iso_code
WHERE     (m.mutation_id NOT LIKE '%none%') AND (p.age_of_onset <> ' ')
ORDER BY 'Gene', m.codon_id
```

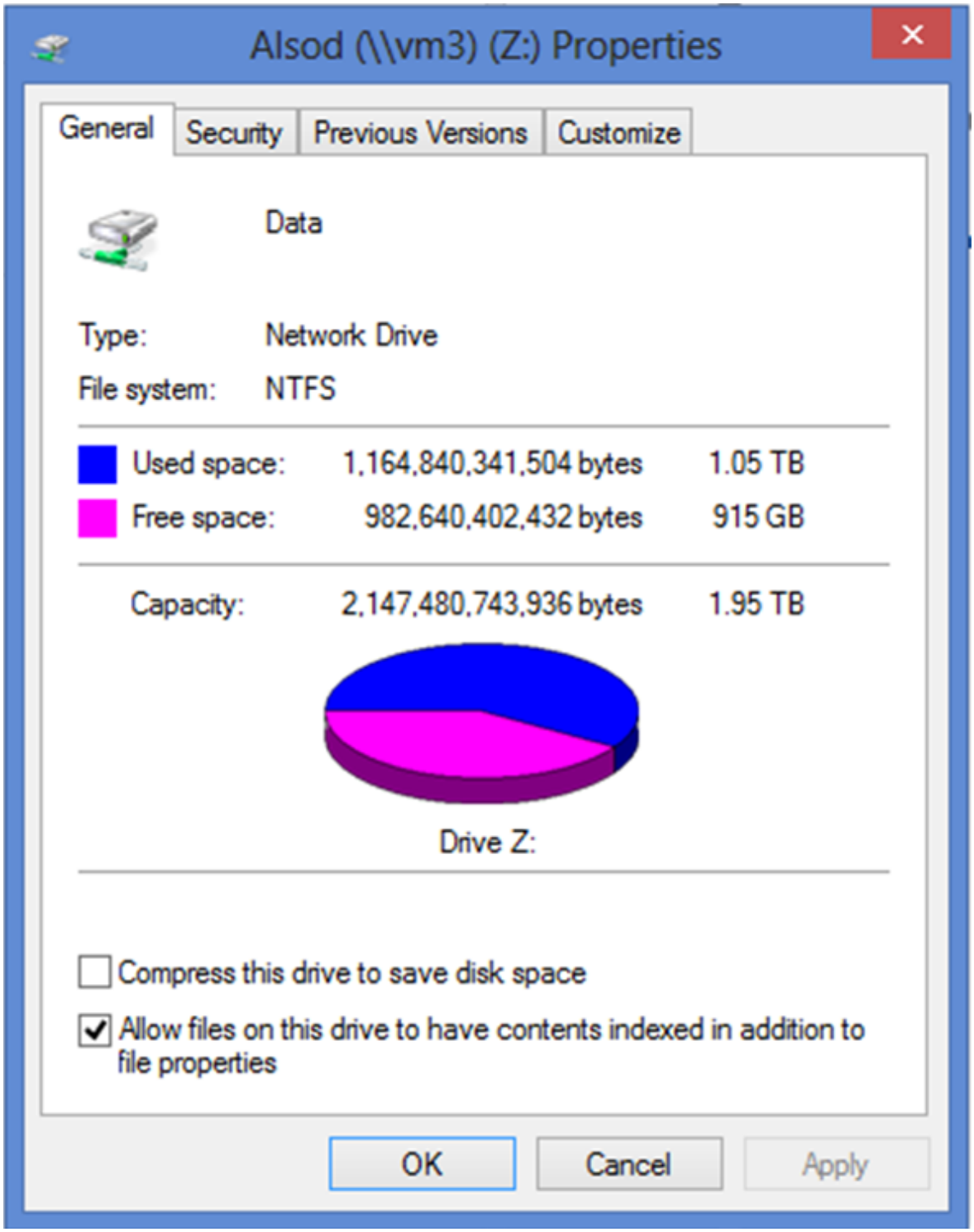
Polyphen_Sift

```
SELECT     dbo.POLYPHEN_scores.gene, dbo.POLYPHEN_scores.mutation,
dbo.POLYPHEN_scores.codon, dbo.POLYPHEN_scores.PSIC_score AS POLYPHEN,
                      dbo.POLYPHEN_scores.prediction AS Prediction2,
dbo.SIFT_scores.subPSEC AS SIFT, dbo.SIFT_scores.message AS Prediction3
FROM         dbo.POLYPHEN_scores INNER JOIN
                      dbo.SIFT_scores ON dbo.POLYPHEN_scores.gene =
dbo.SIFT_scores.gene AND dbo.POLYPHEN_scores.codon = dbo.SIFT_scores.codon
WHERE     (dbo.POLYPHEN_scores.PSIC_score IS NOT NULL) AND
(dbo.SIFT_scores.subPSEC IS NOT NULL)
```

Ranking_LOD_score

```
SELECT DISTINCT gene, LODscore, author, year, DENSE_RANK() OVER (ORDER BY
LODscore DESC) AS Rank FROM dbo.mutation_replication WHERE  (LODscore IS
NOT NULL) OR (LODscore <> '0')
```

Appendix 26 – Disk Usage by Table in ALSoD



Microsoft SQL Server 2008

Disk Usage by Table

[Alsod]

on VM3 at 13/02/2014 16:27:25

This report provides detailed data on the utilization of disk space by tables within the Database.

Table Name	# Records	Reserved (KB)	Data (KB)	Indexes (KB)	Unused (KB)
dbo.amino_acid	22	24	8	16	0
dbo.ANALYZE_JOINTDATA	4	16	8	8	0
dbo.ANALYZE_NOKEY	1,420,354	68,824	68,752	24	48
dbo.ANALYZE_OURDATA	288,136	40,264	40,072	144	48
dbo.ANALYZE_OURDATA_OLD	281,907	23,504	23,280	160	64
dbo.ANALYZE_THEIRDATA	4	16	8	8	0
dbo.Animal_Pubmed	2,404	392	384	8	0
dbo.AnimalModels	4,404	1,680	1,440	8	232
dbo.aspnet_Applications	2	64	8	56	0
dbo.aspnet_Membership	111	176	72	56	48
dbo.aspnet_Paths	0	0	0	0	0
dbo.aspnet_PersonalizationAllUsers	0	0	0	0	0
dbo.aspnet_PersonalizationPerUser	0	0	0	0	0
dbo.aspnet_Profile	0	0	0	0	0
dbo.aspnet_Roles	0	0	0	0	0
dbo.aspnet_SchemaVersions	6	16	8	8	0
dbo.aspnet_Users	111	64	16	48	0
dbo.aspnet_UsersInRoles	0	0	0	0	0
dbo.aspnet_WebEvent_Events	0	0	0	0	0
dbo.codons	7,414	984	456	200	328
dbo.codons1	414	40	32	8	0
dbo.codons2	147	16	8	8	0
dbo.codons3	526	40	32	8	0
dbo.codons4	2,677	264	152	8	104
dbo.codons5	577	40	32	8	0
dbo.Continent	7	24	8	16	0
dbo.counter	1	16	8	8	0
dbo.Country	234	40	8	32	0
dbo.countrydetails	236	56	40	16	0
dbo.countryinfo	244	48	40	8	0
dbo.countrylatlong	241	16	8	8	0
dbo.credibility_survey	5	16	8	8	0
dbo.dtproperties	35	288	256	16	16
dbo.Epidemiology	21	16	8	8	0
dbo.gene	112	624	384	16	224
dbo.gene_authors	201	96	80	16	0
dbo.gene_frequency	1	16	8	8	0
dbo.gene_sequence	325,001	20,640	15,456	5,128	56
dbo.gene_sequence1	13,871	1,032	984	8	40
dbo.gene_sequence2	11,010	712	672	8	32
dbo.gene_sequence3	12,647	840	776	8	56
dbo.gene_sequence4	94,546	5,896	5,848	8	40
dbo.gene_sequence5	39,195	2,440	2,432	8	0
dbo.gene_SNP	47	40	32	8	0
dbo.gene_SNP_ALSGene	413	136	96	8	32
dbo.gene_SNP_lifeseq	2,752	1,304	1,216	8	80
dbo.gene_SNP_updated	1,422	520	352	8	160
dbo.gene_study	286	544	304	24	216
dbo.guestbook	26	96	88	8	0
dbo.GWA_BOS	282,448	52,304	52,056	184	64
dbo.GWA_CATALOGUE	4,811	2,384	2,336	16	32
dbo.GWA_CATALOGUE_1	4,239	2,144	2,072	32	40
dbo.GWA_CATALOGUE2	0	0	0	0	0
dbo.GWA_FOGH	6,138,741	236,184	234,688	1,456	40
dbo.GWA_FRA	286,507	53,008	52,808	192	8
dbo.GWA_HOL	288,136	53,328	53,104	192	32
dbo.GWA_NIH	287,640	53,264	53,016	192	56
dbo.GWA_NOKEY	1,420,350	150,800	150,760	16	24
dbo.GWA_UK	275,619	51,024	50,792	184	48
dbo.GWA_UK_BOS	423,672	41,936	41,744	128	64
dbo.GWA_UK_BOS_FRA	286,507	42,832	42,712	112	8
dbo.GWA_UK_BOS_FRA_NIH	287,640	70,984	70,760	192	32
dbo.GWA_UK_BOS_FRA_USA	0	0	0	0	0
dbo.GWA_UK_BOS_HOL_FRA_NIH	288,136	118,480	118,216	248	16
dbo.GWA_UK_BOS_HOL_FRA_USA	288,136	118,552	118,216	320	16
dbo.GWA_USA	287,640	33,944	33,680	200	64
dbo.GWA3_BEL	224,629	45,256	45,048	160	48
dbo.GWA3_FRA	224,629	45,272	45,048	176	48
dbo.GWA3_HOL	224,629	45,256	45,048	160	48
dbo.GWA3_IRE	224,629	22,728	22,528	144	56
dbo.GWA3_ITA	224,629	45,256	45,048	160	48
dbo.GWA3_NOKEY	1,797,032	206,296	206,232	24	40
dbo.GWA3_POPULATIONS	255	24	16	8	0
dbo.GWA3_SWE	224,629	45,256	45,048	160	48
dbo.GWA3_UK	224,629	45,256	45,048	160	48
dbo.GWA3_USA	224,629	45,272	45,048	176	48
dbo.institution	91	48	32	16	0
dbo.institution1	57	64	32	32	0
dbo.ipcountry	95,296	5,960	5,952	8	0
dbo.ipcountry2	0	0	0	0	0
dbo.ipcountry3	65,536	3,280	3,232	16	32
dbo.members	0	0	0	0	0
dbo.mutation	531	1,232	864	16	352
dbo.mutation_new	0	32	24	8	0
dbo.mutation_replication	811	456	312	8	136
dbo.mutation_replication2	860	200	112	8	80
dbo.mutation_replication3	541	136	88	8	40
dbo.mutation1	137	152	96	32	24
dbo.NETPHOS_scores	1,295	136	72	8	56
dbo.No_PATHOGENICITY	1	24	16	8	0
dbo.No_SALS	4	16	8	8	0
dbo.PANTHER_scores	146,866	20,424	20,360	8	56
dbo.patient_clinical_record	853	200	120	16	64
dbo.patient_genetic_record	929	520	344	16	160
dbo.patient_genetic_record1	303	136	104	16	16
dbo.patient_genetic_record2	168	136	64	16	56
dbo.patient_record	958	616	392	24	200
dbo.POLYPHEN_scores	428	56	48	8	0
dbo.population_frequency_1000genome	56	16	8	8	0
dbo.population_frequency_EVS	89	24	16	8	0
dbo.random_subpsec	1,008	56	40	16	0
dbo.Reference_MGI	557,280	22,792	22,672	56	64
dbo.Reference_Pubmed	27,894	3,848	3,760	16	72
dbo.refs	71	64	32	32	0
dbo.refs_lookup	176	24	16	8	0
dbo.SIFT_scores	12,276	712	704	8	0
dbo.smadmin4	1	40	8	32	0
dbo.smagent_agentgroup4	0	0	0	0	0
dbo.smagent4	1	56	8	48	0
dbo.smagentcommand4	0	0	0	0	0
dbo.smagentgroup4	0	0	0	0	0
dbo.smagentkey4	4	24	8	16	0
dbo.smagenttype4	14	88	8	80	0
dbo.smagenttypeactions4	22	24	8	16	0
dbo.smagenttypeattribute4	158	104	40	64	0
dbo.smagenttypeattributevalues4	172	48	32	16	0

Appendix 27 – Label visibility on credibility analysis webpage

```
<asp:Label ID="Label1" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label2" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label3" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label4" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label5" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label6" runat="server" Visible="false"></asp:Label><br />

 <asp:Label ID="Labela" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Labelb" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Labelc" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Labeld" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Labele" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Labelf" runat="server" Visible="false"></asp:Label><br />

 <asp:Label ID="Label2a" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label2b" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label2c" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label2d" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label2e" runat="server" Visible="false"></asp:Label><br />
 <asp:Label ID="Label2f" runat="server" Visible="false"></asp:Label><br />
```

To cross check the scripts generated from the formula during execution and debugging, the label visibility was set to ‘true’.

```
<asp:Label ID="Label1" runat="server" Visible="true"></asp:Label><br />
   <asp:Label ID="Label2" runat="server" Visible="true"></asp:Label><br />
   <asp:Label ID="Label3" runat="server" Visible="true"></asp:Label><br />
   <asp:Label ID="Label4" runat="server" Visible="true"></asp:Label><br />
   <asp:Label ID="Label5" runat="server" Visible="true"></asp:Label><br />
   <asp:Label ID="Label6" runat="server" Visible="true"></asp:Label><br />

   <asp:Label ID="Labela" runat="server" Visible="true"></asp:Label><br />
```

```
<asp:Label ID="Labelb" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Labelc" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Labeld" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Labele" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Labelf" runat="server" Visible="true"></asp:Label><br />

<asp:Label ID="Label2a" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Label2b" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Label2c" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Label2d" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Label2e" runat="server" Visible="true"></asp:Label><br />
<asp:Label ID="Label2f" runat="server" Visible="true"></asp:Label><br />
```

Appendix 28 – User interface ASP.NET code

```
<asp:CheckBoxList ID="tables" runat="server" AutoPostBack ="false"

                ToolTip="Select checkbox if desired as one of the criteria for analysing credibility score" RepeatColumns="3"

                RepeatDirection="Horizontal">

            <asp:ListItem       Value="[KCLAD\spngoka].[Number_of_affected_patients_in_ALSoD]" Text="Rank_Patients" Enabled="false" ></asp:ListItem>

            <asp:ListItem               Value="[KCLAD\spngoka].[Number_of_mutations_per_gene]" Text="Rank_Mutations" Enabled="false"></asp:ListItem>

            <asp:ListItem                   Value="[KCLAD\spngoka].[Number_of_cases_recorded]" Text="Rank_Cases"></asp:ListItem>

            <asp:ListItem                   Value="[KCLAD\spngoka].[Number_of_controls_recorded]" Text="Rank_Controls"></asp:ListItem>

            <asp:ListItem                       Value="[KCLAD\spngoka].[Number_of_mutations_in_same_codon_by_rank]" Text="Rank_Codon"></asp:ListItem>

            <asp:ListItem
Value="[KCLAD\spngoka].[Number_of_patients_with_family_history_FALS]"
Text="Rank_FALS"></asp:ListItem>

            <asp:ListItem
Value="[KCLAD\spngoka].[Number_of_patients_without_family_history_SALS]"
Text="Rank_SALS"></asp:ListItem>

            <asp:ListItem       Value="[KCLAD\spngoka].[Number_of_times_mutation_is_replicated]" Text="Rank_Replications"></asp:ListItem>

            <asp:ListItem    Value="[KCLAD\spngoka].[Number_of_pathogenic_mutations_by_rank]" Text="Rank_Pathogenicity"></asp:ListItem>

            <asp:ListItem       Value="[KCLAD\spngoka].[Number_of_unique_countries_on_genes]" Text="Rank_Populations"></asp:ListItem>


        </asp:CheckBoxList>
```

Appendix 29 – MainActivity.Java for Android app

```
import android.os.Bundle;
import android.annotation.SuppressLint;
import android.app.Activity;
import android.view.Menu;
import android.webkit.WebView;

public class MainActivity extends Activity {
        WebView browser;

    @SuppressLint("SetJavaScriptEnabled")
        @Override
    public void onCreate(Bundle savedInstanceState) {
      super.onCreate(savedInstanceState);
      setContentView(R.layout.activity_main);

     // find the WebView by name in the main.xml of step 2
      browser=(WebView)findViewById(R.id.wvwMain);

      // Enable javascript
      browser.getSettings().setJavaScriptEnabled(true);

      // load a webpage
      browser.loadUrl("http://alsod.iop.kcl.ac.uk/Mobile/index.aspx");

    }
    @Override
    public boolean onCreateOptionsMenu(Menu menu) {
      getMenuInflater().inflate(R.menu.activity_main, menu);
      return true;
    }
}
```

Appendix 30 – Possible graphs of 31 combinations for 5 populations

```
If Label1.Text = "UK" Then
    Image4.ImageUrl = "~/PlotGraphs/Uk.png"
ElseIf Label1.Text = "BOSTON" Then
    Image4.ImageUrl = "~/PlotGraphs/Bos.png"
ElseIf Label1.Text = "HOLLAND" Then
    Image4.ImageUrl = "~/PlotGraphs/Hol.png"
ElseIf Label1.Text = "FRANCE" Then
    Image4.ImageUrl = "~/PlotGraphs/Fra.png"
ElseIf Label1.Text = "NIH" Then
    Image4.ImageUrl = "~/PlotGraphs/Usa.png"
ElseIf Label1.Text = "UKBOSTON" Then
    Image4.ImageUrl = "~/PlotGraphs/UkBos.png"
ElseIf Label1.Text = "UKHOLLAND" Then
    Image4.ImageUrl = "~/PlotGraphs/UkHol.png"
ElseIf Label1.Text = "UKFRANCE" Then
    Image4.ImageUrl = "~/PlotGraphs/UkFra.png"
ElseIf Label1.Text = "UKNIH" Then
    Image4.ImageUrl = "~/PlotGraphs/UkUsa.png"
ElseIf Label1.Text = "UKBOSTONHOLLAND" Then
    Image4.ImageUrl = "~/PlotGraphs/UkBosHol.png"
```

```
ElseIf Label1.Text = "UKBOSTONFRANCE" Then
    Image4.ImageUrl = "~/PlotGraphs/UkBosFra.png"
ElseIf Label1.Text = "UKBOSTONNIH" Then
    Image4.ImageUrl = "~/PlotGraphs/UkBosUsa.png"
ElseIf Label1.Text = "UKHOLLANDFRANCE" Then
    Image4.ImageUrl = "~/PlotGraphs/UkHolFra.png"
ElseIf Label1.Text = "UKHOLLANDNIH" Then
    Image4.ImageUrl = "~/PlotGraphs/UkHolUsa.png"
ElseIf Label1.Text = "UKFRANCENIH" Then
    Image4.ImageUrl = "~/PlotGraphs/UkFraUsa.png"
ElseIf Label1.Text = "UKBOSTONHOLLANDFRANCE" Then
    Image4.ImageUrl = "~/PlotGraphs/UkBosHolFra.png"
ElseIf Label1.Text = "UKBOSTONHOLLANDNIH" Then
    Image4.ImageUrl = "~/PlotGraphs/UkBosHolUsa.png"
ElseIf Label1.Text = "UKBOSTONFRANCENIH" Then
    Image4.ImageUrl = "~/PlotGraphs/UkBosFraUsa.png"
ElseIf Label1.Text = "UKHOLLANDFRANCENIH" Then
    Image4.ImageUrl = "~/PlotGraphs/UkHolFraUsa.png"
ElseIf Label1.Text = "BOSTONHOLLAND" Then
    Image4.ImageUrl = "~/PlotGraphs/BosHol.png"
ElseIf Label1.Text = "BOSTONFRANCE" Then
    Image4.ImageUrl = "~/PlotGraphs/BosFra.png"
```

```
ElseIf Label1.Text = "BOSTONNIH" Then
    Image4.ImageUrl = "~/PlotGraphs/BosUsa.png"
ElseIf Label1.Text = "BOSTONHOLLANDFRANCE" Then
    Image4.ImageUrl = "~/PlotGraphs/BosHolFra.png"
ElseIf Label1.Text = "BOSTONHOLLANDNIH" Then
    Image4.ImageUrl = "~/PlotGraphs/BosHolUsa.png"
ElseIf Label1.Text = "BOSTONFRANCENIH" Then
    Image4.ImageUrl = "~/PlotGraphs/BosFraUsa.png"
ElseIf Label1.Text = "BOSTONHOLLANDFRANCENIH" Then
    Image4.ImageUrl = "~/PlotGraphs/BosHolFraUsa.png"
ElseIf Label1.Text = "HOLLANDFRANCE" Then
    Image4.ImageUrl = "~/PlotGraphs/HolFra.png"
ElseIf Label1.Text = "HOLLANDNIH" Then
    Image4.ImageUrl = "~/PlotGraphs/HolUsa.png"
ElseIf Label1.Text = "HOLLANDFRANCENIH" Then
    Image4.ImageUrl = "~/PlotGraphs/HolFraUsa.png"
ElseIf Label1.Text = "FRANCENIH" Then
    Image4.ImageUrl = "~/PlotGraphs/FraUsa.png"
ElseIf Label1.Text = "UKBOSTONHOLLANDFRANCENIH" Then
    Image4.ImageUrl = "~/PlotGraphs/UkBosHolFraUsa.png"
Else : table3.Visible = "False"
    table2.Visible = "True"
```

```
End If
```

Appendix 31 – UCSC Genome Browser

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2//EN">

<HTML><HEAD>

<TITLE>Human chrX:151,073,054-151,383,976 - UCSC Genome Browser v187</TITLE>

<META http-equiv="Content-Script-Type" content="text/javascript">

<LINK REL="STYLESHEET" HREF="../style/HGStyle.css" TYPE="text/css">

</HEAD>

<BODY BACKGROUND="../images/floret.jpg">

<FORM ACTION="../cgi-bin/hgTracks" NAME="TrackHeaderForm" METHOD=GET>

<INPUT TYPE=HIDDEN NAME="hgsid" VALUE="109992416"><CENTER>

<DIV STYLE="white-space:nowrap;">

<TABLE WIDTH="100%" BGCOLOR="#000000" BORDER="0" CELLSPACING="0" CELLPADDING="1"><TR><TD>

<TABLE WIDTH="100%" BGCOLOR="#2636D1" BORDER="0" CELLSPACING="0" CELLPADDING="2"><TR>

<TD ALIGN=CENTER><A HREF="../index.html?org=Human&db=hg18&hgsid=109992416" class="topbar">Home</A></TD><TD ALIGN=CENTER><A HREF="../cgi-bin/hgGateway?org=Human&db=hg18&hgsid=109992416" class="topbar">Genomes</A></TD><TD ALIGN=CENTER><A HREF="../cgi-bin/hgBlat?hgsid=109992416" class="topbar">Blat</A></TD><TD ALIGN=CENTER><A HREF="../cgi-bin/hgTables?db=hg18&position=chrX:151073054-151383976&hgsid=109992416" class="topbar">Tables</A></TD><TD ALIGN=CENTER><A HREF="../cgi-bin/hgNear?hgsid=109992416" class="topbar">Gene Sorter</A></TD><TD ALIGN=CENTER><A HREF="../cgi-bin/hgPcr?hgsid=109992416" class="topbar">PCR</A></TD><TD ALIGN=CENTER><A HREF="../cgi-bin/hgc?hgsid=109992416&o=151073053&g=getDna&i=mixed&c=chrX&l=151073053&r=151383976&db=hg18&hgsid=109992416" class="topbar"> DNA </A></TD><TD ALIGN=CENTER><A HREF="../cgi-bin/hgConvert?hgsid=109992416&db=hg18&position=chrX:151073054-151383976" class="topbar">Convert</A></TD><TD ALIGN=CENTER>

<A HREF="http://www.ensembl.org/Homo_sapiens/contigview?chr=X&start=151073054&end=151383976" TARGET=_blank class="topbar">Ensembl</A></TD><TD ALIGN=CENTER><A HREF="http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?taxid=9606&CHR=X&BEG=151073054&END=151383976" TARGET=_blank class="topbar">NCBI</A></TD><TD ALIGN=CENTER><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgt.psOutput=on" class="topbar">PDF/PS</A></TD>

<TD ALIGN=CENTER><A HREF="../cgi-bin/hgSession?hgsid=109992416&hgS_doMainPage=1" class="topbar">Session</A></TD><TD ALIGN=CENTER><A HREF="../goldenPath/help/hgTracksHelp.html" TARGET=_blank class="topbar">Help</A></TD>

</TR></TABLE>

</TD></TR></TABLE>

<FONT SIZE=5><B>UCSC Genome Browser on Human Mar. 2006 Assembly</B></FONT>

<INPUT TYPE=IMAGE BORDER=0 NAME="hgt.dummyEnterButton" src="../images/DOT.gif">move <INPUT TYPE=SUBMIT NAME="hgt.left3" VALUE="<<<" ><INPUT TYPE=SUBMIT NAME="hgt.left2" VALUE=" <<" ><INPUT TYPE=SUBMIT NAME="hgt.left1" VALUE=" < " ><INPUT TYPE=SUBMIT NAME="hgt.right1" VALUE=" > " ><INPUT TYPE=SUBMIT NAME="hgt.right2" VALUE=">> " ><INPUT TYPE=SUBMIT NAME="hgt.right3" VALUE=">>>" > zoom in <INPUT TYPE=SUBMIT NAME="hgt.in1" VALUE="1.5x" ><INPUT TYPE=SUBMIT NAME="hgt.in2" VALUE=" 3x " ><INPUT TYPE=SUBMIT NAME="hgt.in3" VALUE="10x" ><INPUT TYPE=SUBMIT NAME="hgt.inBase" VALUE="base" > zoom out <INPUT TYPE=SUBMIT NAME="hgt.out1" VALUE="1.5x" ><INPUT TYPE=SUBMIT NAME="hgt.out2" VALUE=" 3x " ><INPUT TYPE=SUBMIT NAME="hgt.out3" VALUE="10x" >

<INPUT TYPE=HIDDEN NAME="position" VALUE="chrX:151073054-151383976">

<input type='hidden' name="hgtgroup_map_close" id="hgtgroup_map_close_1" value="0">

<input type='hidden' name="hgtgroup_phenDis_close" id="hgtgroup_phenDis_close_1" value="0">

<input type='hidden' name="hgtgroup_genes_close" id="hgtgroup_genes_close_1" value="0">

<input type='hidden' name="hgtgroup_rna_close" id="hgtgroup_rna_close_1" value="0">

<input type='hidden' name="hgtgroup_regulation_close" id="hgtgroup_regulation_close_1" value="0">

<input type='hidden' name="hgtgroup_compGeno_close" id="hgtgroup_compGeno_close_1" value="0">

<input type='hidden' name="hgtgroup_varRep_close" id="hgtgroup_varRep_close_1" value="0">

<input type='hidden' name="hgtgroup_encodeGenes_close" id="hgtgroup_encodeGenes_close_1" value="0">

<input type='hidden' name="hgtgroup_encodeTxLevels_close" id="hgtgroup_encodeTxLevels_close_1" value="0">

<input type='hidden' name="hgtgroup_encodeChip_close" id="hgtgroup_encodeChip_close_1" value="0">

<input type='hidden' name="hgtgroup_encodeChrom_close" id="hgtgroup_encodeChrom_close_1" value="0">

<input type='hidden' name="hgtgroup_encodeCompAndVar_close" id="hgtgroup_encodeCompAndVar_close_1" value="0">

</CENTER></FORM>

<FORM ACTION="../cgi-bin/hgTracks" NAME="TrackForm" METHOD=POST>

<input type='hidden' name="hgtgroup_map_close" id="hgtgroup_map_close_2" value="0">

<input type='hidden' name="hgtgroup_phenDis_close" id="hgtgroup_phenDis_close_2" value="0">

<input type='hidden' name="hgtgroup_genes_close" id="hgtgroup_genes_close_2" value="0">

<input type='hidden' name="hgtgroup_rna_close" id="hgtgroup_rna_close_2" value="0">

<input type='hidden' name="hgtgroup_regulation_close" id="hgtgroup_regulation_close_2" value="0">

<input type='hidden' name="hgtgroup_compGeno_close" id="hgtgroup_compGeno_close_2" value="0">

<input type='hidden' name="hgtgroup_varRep_close" id="hgtgroup_varRep_close_2" value="0">

<input type='hidden' name="hgtgroup_encodeGenes_close" id="hgtgroup_encodeGenes_close_2" value="0">

<input type='hidden' name="hgtgroup_encodeTxLevels_close" id="hgtgroup_encodeTxLevels_close_2" value="0">

<input type='hidden' name="hgtgroup_encodeChip_close" id="hgtgroup_encodeChip_close_2" value="0">

<input type='hidden' name="hgtgroup_encodeChrom_close" id="hgtgroup_encodeChrom_close_2" value="0">

<input type='hidden' name="hgtgroup_encodeCompAndVar_close" id="hgtgroup_encodeCompAndVar_close_2" value="0">

<INPUT TYPE=HIDDEN NAME="hgsid" VALUE="109992416"><CENTER>position/search <INPUT TYPE=TEXT NAME="position" SIZE=30 VALUE="chrX:151,073,054-151,383,976">

<INPUT TYPE=SUBMIT NAME="submit" VALUE="jump" ><INPUT TYPE="button" VALUE="clear" onClick="document.TrackForm.position.value=""> size 310,923 bp. <INPUT TYPE=SUBMIT NAME="hgTracksConfigPage" VALUE="configure" >

<MAP Name=ideoMap>

<AREA SHAPE=RECT COORDS="68,6,80,15" HREF="../cgi-bin/hgTracks?position=chrX:1-4300000&hgsid=109992416" TITLE="p22.33">

<AREA SHAPE=RECT COORDS="80,6,84,15" HREF="../cgi-bin/hgTracks?position=chrX:4300001-6000000&hgsid=109992416" TITLE="p22.32">

<AREA SHAPE=RECT COORDS="84,6,94,15" HREF="../cgi-bin/hgTracks?position=chrX:6000001-9500000&hgsid=109992416" TITLE="p22.31">

<AREA SHAPE=RECT COORDS="94,6,115,15" HREF="../cgi-bin/hgTracks?position=chrX:9500001-17100000&hgsid=109992416" TITLE="p22.2">

<AREA SHAPE=RECT COORDS="115,6,121,15" HREF="../cgi-bin/hgTracks?position=chrX:17100001-19200000&hgsid=109992416" TITLE="p22.13">

<AREA SHAPE=RECT COORDS="121,6,128,15" HREF="../cgi-bin/hgTracks?position=chrX:19200001-21800000&hgsid=109992416" TITLE="p22.12">

<AREA SHAPE=RECT COORDS="128,6,136,15" HREF="../cgi-bin/hgTracks?position=chrX:21800001-24900000&hgsid=109992416" TITLE="p22.11">

<AREA SHAPE=RECT COORDS="136,6,148,15" HREF="../cgi-bin/hgTracks?position=chrX:24900001-29400000&hgsid=109992416" TITLE="p21.3">

<AREA SHAPE=RECT COORDS="148,6,154,15" HREF="../cgi-bin/hgTracks?position=chrX:29400001-31500000&hgsid=109992416" TITLE="p21.2">

<AREA SHAPE=RECT COORDS="154,6,171,15" HREF="../cgi-bin/hgTracks?position=chrX:31500001-37500000&hgsid=109992416" TITLE="p21.1">

<AREA SHAPE=RECT COORDS="171,6,184,15" HREF="../cgi-bin/hgTracks?position=chrX:37500001-42300000&hgsid=109992416" TITLE="p11.4">

<AREA SHAPE=RECT COORDS="184,6,197,15" HREF="../cgi-bin/hgTracks?position=chrX:42300001-47300000&hgsid=109992416" TITLE="p11.3">

<AREA SHAPE=RECT COORDS="197,6,204,15" HREF="../cgi-bin/hgTracks?position=chrX:47300001-49700000&hgsid=109992416" TITLE="p11.23">

<AREA SHAPE=RECT COORDS="204,6,218,15" HREF="../cgi-bin/hgTracks?position=chrX:49700001-54700000&hgsid=109992416" TITLE="p11.22">

<AREA SHAPE=RECT COORDS="218,6,223,15" HREF="../cgi-bin/hgTracks?position=chrX:54700001-56600000&hgsid=109992416" TITLE="p11.21">

<AREA SHAPE=RECT COORDS="223,6,231,15" HREF="../cgi-bin/hgTracks?position=chrX:56600001-59500000&hgsid=109992416" TITLE="p11.1">

<AREA SHAPE=RECT COORDS="231,6,246,15" HREF="../cgi-bin/hgTracks?position=chrX:59500001-65000000&hgsid=109992416" TITLE="q11.1">

<AREA SHAPE=RECT COORDS="246,6,247,15" HREF="../cgi-bin/hgTracks?position=chrX:65000001-65100000&hgsid=109992416" TITLE="q11.2">

<AREA SHAPE=RECT COORDS="246,6,253,15" HREF="../cgi-bin/hgTracks?position=chrX:65100001-67700000&hgsid=109992416" TITLE="q12">

<AREA SHAPE=RECT COORDS="253,6,266,15" HREF="../cgi-bin/hgTracks?position=chrX:67700001-72200000&hgsid=109992416" TITLE="q13.1">

<AREA SHAPE=RECT COORDS="266,6,270,15" HREF="../cgi-bin/hgTracks?position=chrX:72200001-73800000&hgsid=109992416" TITLE="q13.2">

<AREA SHAPE=RECT COORDS="270,6,276,15" HREF="../cgi-bin/hgTracks?position=chrX:73800001-76000000&hgsid=109992416" TITLE="q13.3">

<AREA SHAPE=RECT COORDS="276,6,299,15" HREF="../cgi-bin/hgTracks?position=chrX:76000001-84500000&hgsid=109992416" TITLE="q21.1">

<AREA SHAPE=RECT COORDS="299,6,304,15" HREF="../cgi-bin/hgTracks?position=chrX:84500001-86200000&hgsid=109992416" TITLE="q21.2">

<AREA SHAPE=RECT COORDS="304,6,320,15" HREF="../cgi-bin/hgTracks?position=chrX:86200001-91900000&hgsid=109992416" TITLE="q21.31">

<AREA SHAPE=RECT COORDS="320,6,324,15" HREF="../cgi-bin/hgTracks?position=chrX:91900001-93500000&hgsid=109992416" TITLE="q21.32">

<AREA SHAPE=RECT COORDS="324,6,337,15" HREF="../cgi-bin/hgTracks?position=chrX:93500001-98200000&hgsid=109992416" TITLE="q21.33">

<AREA SHAPE=RECT COORDS="337,6,349,15" HREF="../cgi-bin/hgTracks?position=chrX:98200001-102500000&hgsid=109992416" TITLE="q22.1">

<AREA SHAPE=RECT COORDS="349,6,352,15" HREF="../cgi-bin/hgTracks?position=chrX:102500001-103600000&hgsid=109992416" TITLE="q22.2">

<AREA SHAPE=RECT COORDS="352,6,370,15" HREF="../cgi-bin/hgTracks?position=chrX:103600001-110500000&hgsid=109992416" TITLE="q22.3">

<AREA SHAPE=RECT COORDS="370,6,388,15" HREF="../cgi-bin/hgTracks?position=chrX:110500001-116800000&hgsid=109992416" TITLE="q23">

<AREA SHAPE=RECT COORDS="388,6,398,15" HREF="../cgi-bin/hgTracks?position=chrX:116800001-120700000&hgsid=109992416" TITLE="q24">

<AREA SHAPE=RECT COORDS="398,6,423,15" HREF="../cgi-bin/hgTracks?position=chrX:120700001-129800000&hgsid=109992416" TITLE="q25">

<AREA SHAPE=RECT COORDS="423,6,425,15" HREF="../cgi-bin/hgTracks?position=chrX:129800001-130300000&hgsid=109992416" TITLE="q26.1">

<AREA SHAPE=RECT COORDS="425,6,433,15" HREF="../cgi-bin/hgTracks?position=chrX:130300001-133500000&hgsid=109992416" TITLE="q26.2">

<AREA SHAPE=RECT COORDS="433,6,445,15" HREF="../cgi-bin/hgTracks?position=chrX:133500001-137800000&hgsid=109992416" TITLE="q26.3">

<AREA SHAPE=RECT COORDS="445,6,451,15" HREF="../cgi-bin/hgTracks?position=chrX:137800001-140100000&hgsid=109992416" TITLE="q27.1">

<AREA SHAPE=RECT COORDS="451,6,456,15" HREF="../cgi-bin/hgTracks?position=chrX:140100001-141900000&hgsid=109992416" TITLE="q27.2">

<AREA SHAPE=RECT COORDS="456,6,470,15" HREF="../cgi-bin/hgTracks?position=chrX:141900001-146900000&hgsid=109992416" TITLE="q27.3">

<AREA SHAPE=RECT COORDS="470,6,492,15" HREF="../cgi-bin/hgTracks?position=chrX:146900001-154913754&hgsid=109992416" TITLE="q28">

</MAP>

<TABLE BORDER=0 CELLPADDING=0><TR><TD HEIGHT=5></TD></TR><TR><TD><IMG SRC = "../trash/hgtIdeo/hgtIdeo_genome_5c4e_4392c0.gif" BORDER=1 WIDTH=496 HEIGHT=21 USEMAP=#ideoMap ></TD></TR><TR><TD HEIGHT=5></TD></TR></TABLE><MAP Name=map>

<AREA SHAPE=RECT COORDS="1,1,12,10" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=ruler" TITLE="Base Position controls">

<AREA SHAPE=RECT COORDS="1,11,12,29" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=knownGene" TITLE="UCSC Genes controls">

<AREA SHAPE=RECT COORDS="1,30,12,48" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=refGene" TITLE="RefSeq Genes controls">

<AREA SHAPE=RECT COORDS="1,49,12,67" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=mgcGenes" TITLE="MGC Genes controls">

<AREA SHAPE=RECT COORDS="1,68,12,86" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=mrna" TITLE="Human mRNAs controls">

<AREA SHAPE=RECT COORDS="1,87,12,105" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=intronEst" TITLE="Spliced ESTs controls">

<AREA SHAPE=RECT COORDS="1,106,12,254" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=multiz28way" TITLE="Conservation controls">

<AREA SHAPE=RECT COORDS="1,255,12,273" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=snp129" TITLE="SNPs (129) controls">

<AREA SHAPE=RECT COORDS="1,274,12,292" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=snp128" TITLE="SNPs (128) controls">

<AREA SHAPE=RECT COORDS="1,293,12,311" HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=rmsk" TITLE="RepeatMasker controls">

<AREA SHAPE=RECT COORDS="115,0,132,9" HREF="../cgi-bin/hgTracks?position=chrX:151026416-151130056&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="132,0,149,9" HREF="../cgi-bin/hgTracks?position=chrX:151036780-151140420&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="149,0,165,9" HREF="../cgi-bin/hgTracks?position=chrX:151047144-151150784&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="165,0,182,9" HREF="../cgi-bin/hgTracks?position=chrX:151057508-151161148&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="182,0,199,9" HREF="../cgi-bin/hgTracks?position=chrX:151067872-151171512&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="199,0,216,9" HREF="../cgi-bin/hgTracks?position=chrX:151078237-151181877&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="216,0,233,9" HREF="../cgi-bin/hgTracks?position=chrX:151088601-151192241&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="233,0,249,9" HREF="../cgi-bin/hgTracks?position=chrX:151098965-151202605&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="249,0,266,9" HREF="../cgi-bin/hgTracks?position=chrX:151109329-151212969&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="266,0,283,9" HREF="../cgi-bin/hgTracks?position=chrX:151119693-151223333&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="283,0,300,9" HREF="../cgi-bin/hgTracks?position=chrX:151130057-151233697&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="300,0,317,9" HREF="../cgi-bin/hgTracks?position=chrX:151140421-151244061&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="317,0,333,9" HREF="../cgi-bin/hgTracks?position=chrX:151150785-151254425&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="333,0,350,9" HREF="../cgi-bin/hgTracks?position=chrX:151161149-151264789&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="350,0,367,9" HREF="../cgi-bin/hgTracks?position=chrX:151171513-151275153&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="367,0,384,9" HREF="../cgi-bin/hgTracks?position=chrX:151181878-151285518&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="384,0,401,9" HREF="../cgi-bin/hgTracks?position=chrX:151192242-151295882&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="401,0,417,9" HREF="../cgi-bin/hgTracks?position=chrX:151202606-151306246&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="417,0,434,9" HREF="../cgi-bin/hgTracks?position=chrX:151212970-151316610&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="434,0,451,9" HREF="../cgi-bin/hgTracks?position=chrX:151223334-151326974&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="451,0,468,9" HREF="../cgi-bin/hgTracks?position=chrX:151233698-151337338&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="468,0,485,9" HREF="../cgi-bin/hgTracks?position=chrX:151244062-151347702&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="485,0,501,9" HREF="../cgi-bin/hgTracks?position=chrX:151254426-151358066&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="501,0,518,9" HREF="../cgi-bin/hgTracks?position=chrX:151264790-151368430&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="518,0,535,9" HREF="../cgi-bin/hgTracks?position=chrX:151275154-151378794&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="535,0,552,9" HREF="../cgi-bin/hgTracks?position=chrX:151285519-151389159&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="552,0,569,9" HREF="../cgi-bin/hgTracks?position=chrX:151295883-151399523&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="569,0,585,9" HREF="../cgi-bin/hgTracks?position=chrX:151306247-151409887&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="585,0,602,9" HREF="../cgi-bin/hgTracks?position=chrX:151316611-151420251&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="602,0,619,9" HREF="../cgi-bin/hgTracks?position=chrX:151326975-151430615&hgsid=109992416" TITLE="3x zoom">

<AREA SHAPE=RECT COORDS="11,11,620,19" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&knownGene=pack" TITLE="Toggle the display density of UCSC Genes">

<AREA SHAPE=RECT COORDS="11,11,620,19" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&knownGene=pack" TITLE="Toggle the display density of UCSC Genes">

<AREA SHAPE=RECT COORDS="11,30,620,38" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&refGene=pack" TITLE="Toggle the display density of RefSeq Genes">

<AREA SHAPE=RECT COORDS="11,30,620,38" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&refGene=pack" TITLE="Toggle the display density of RefSeq Genes">

<AREA SHAPE=RECT COORDS="11,49,620,57" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&mgcGenes=dense" TITLE="Toggle the display density of MGC Genes">

<AREA SHAPE=RECT COORDS="11,49,620,57" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&mgcGenes=dense" TITLE="Toggle the display density of MGC Genes">

<AREA SHAPE=RECT COORDS="11,68,620,76" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&mrna=pack" TITLE="Toggle the display density of Human mRNAs">

<AREA SHAPE=RECT COORDS="11,68,620,76" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&mrna=pack" TITLE="Toggle the display density of Human mRNAs">

<AREA SHAPE=RECT COORDS="11,87,620,95" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&intronEst=pack" TITLE="Toggle the display density of Spliced ESTs">

<AREA SHAPE=RECT COORDS="11,87,620,95" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&intronEst=pack" TITLE="Toggle the display density of Spliced ESTs">

<AREA SHAPE=RECT COORDS="11,106,620,114" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&multiz28way=dense" TITLE="Toggle the display density of Conservation">

<AREA SHAPE=RECT COORDS="11,106,620,114" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&multiz28way=dense" TITLE="Toggle the display density of Conservation">

<AREA SHAPE=RECT COORDS="11,255,620,263" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&snp129=pack" TITLE="Toggle the display density of SNPs (129)">

<AREA SHAPE=RECT COORDS="11,255,620,263" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&snp129=pack" TITLE="Toggle the display density of SNPs (129)">

<AREA SHAPE=RECT COORDS="11,274,620,282" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&snp128=pack" TITLE="Toggle the display density of SNPs (128)">

<AREA SHAPE=RECT COORDS="11,274,620,282" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&snp128=pack" TITLE="Toggle the display density of SNPs (128)">

<AREA SHAPE=RECT COORDS="11,293,620,301" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&rmsk=full" TITLE="Toggle the display density of RepeatMasker">

<AREA SHAPE=RECT COORDS="11,293,620,301" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&rmsk=full" TITLE="Toggle the display density of RepeatMasker">

<AREA SHAPE=RECT COORDS="13,57,597,66" HREF="../cgi-bin/hgc?hgsid=109992416&o=151087187&t=151370481&g=mgcGenes&i=BC028629&c=chrX&l=151073053&r=151383976&db=hg18&pix=620" TITLE="BC028629">

<AREA SHAPE=RECT COORDS="115,114,619,254" HREF="../cgi-bin/hgc?hgsid=109992416&o=151073053&t=151383976&g=multiz28way&i=multiz28way&c=chrX&l=151073053&r=151383976&db=hg18&pix=620" >

<AREA SHAPE=RECT COORDS="11,19,620,29" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&knownGene=pack" TITLE="Toggle the display density of UCSC Genes">

<AREA SHAPE=RECT COORDS="11,38,620,48" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&refGene=pack" TITLE="Toggle the display density of RefSeq Genes">

<AREA SHAPE=RECT COORDS="11,76,620,86" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&mrna=pack" TITLE="Toggle the display density of Human mRNAs">

<AREA SHAPE=RECT COORDS="11,95,620,105" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&intronEst=pack" TITLE="Toggle the display density of Spliced ESTs">

<AREA SHAPE=RECT COORDS="11,263,620,273" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&snp129=pack" TITLE="Toggle the display density of SNPs (129)">

<AREA SHAPE=RECT COORDS="11,282,620,292" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&snp128=pack" TITLE="Toggle the display density of SNPs (128)">

<AREA SHAPE=RECT COORDS="11,301,620,311" HREF="../cgi-bin/hgTracks?position=chrX:151073054-151383976&hgsid=109992416&rmsk=full" TITLE="Toggle the display density of RepeatMasker">

</MAP>

<IMG SRC = "../trash/hgt/hgt_genome_5c4e_4392c0.gif" BORDER=1 WIDTH=620 HEIGHT=311 USEMAP=#map >

<TABLE BORDER=0 CELLSPACING=1 CELLPADDING=1 WIDTH=620 COLS=27><TR>

<TD COLSPAN=6 ALIGN=CENTER NOWRAP>move start
<INPUT TYPE=SUBMIT NAME="hgt.dinkLL" VALUE=" < " ><INPUT TYPE=TEXT NAME="dinkL" SIZE=3 VALUE="2.0">

<INPUT TYPE=SUBMIT NAME="hgt.dinkLR" VALUE=" > " ></TD><TD COLSPAN=15 style="white-space:normal">Click on a feature for details. Click on base position to zoom in around cursor. Click gray/blue bars on left for track options and descriptions.</TD><TD COLSPAN=6 ALIGN=CENTER NOWRAP>move end
<INPUT TYPE=SUBMIT NAME="hgt.dinkRL" VALUE=" < " ><INPUT TYPE=TEXT NAME="dinkR" SIZE=3 VALUE="2.0">

<INPUT TYPE=SUBMIT NAME="hgt.dinkRR" VALUE=" > " ></TD></TR></TABLE>

<INPUT TYPE=SUBMIT NAME="hgt.reset" VALUE="default tracks" > <INPUT TYPE=SUBMIT NAME="hgt.hideAll" VALUE="hide all" > <INPUT TYPE="button" VALUE="add custom tracks" onClick="document.customTrackForm.submit();return false;"> <INPUT TYPE=SUBMIT NAME="hgTracksConfigPage" VALUE="configure" > <INPUT TYPE=SUBMIT NAME="hgt.toggleRevCmplDisp" VALUE="reverse" > <INPUT TYPE=SUBMIT NAME="submit" VALUE="refresh" >

<script type='text/javascript' src='../js/hgTracks.js'></script>

<noscript>Your browser does not support JavaScript so some functionality may be missing!</noscript>

<table border=0 cellspacing=1 cellpadding=1 width=610>

<tr><td colspan='5' align='CENTER' nowrap>Use drop-down controls below and press refresh to alter tracks displayed.
Tracks with lots of items will automatically be displayed in more compact modes.</td></tr>

<TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="mapGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_map_close=1#mapGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'map');" id="map_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">Mapping and Sequencing Tracks</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='map-1'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=ruler"> Base Position
 </A><SELECT NAME="ruler" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION SELECTED>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=cytoBand"> Chromosome Band
 </A><SELECT NAME="cytoBand" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=stsMap"> STS Markers
 </A><SELECT NAME="stsMap" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=fishClones"> FISH Clones
 </A><SELECT NAME="fishClones" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=recombRate"> Recomb Rate
 </A><SELECT NAME="recombRate" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='map-2'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=ctgPos"> Map Contigs
 </A><SELECT NAME="ctgPos" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=gold"> Assembly
 </A><SELECT NAME="gold" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=gap"> Gap
 </A><SELECT NAME="gap" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=clonePos"> Coverage
 </A><SELECT NAME="clonePos" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=bacEndPairs"> BAC End Pairs
 </A><SELECT NAME="bacEndPairs" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='map-3'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=fosEndPairs"> Fosmid End Pairs
 </A><SELECT NAME="fosEndPairs" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=gc5Base"> GC Percent
 </A><SELECT NAME="gc5Base" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=oligoMatch"> Short Match
 </A><SELECT NAME="oligoMatch" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=cutters"> Restr Enzymes
 </A><SELECT NAME="cutters" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="phenDisGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_phenDis_close=1#phenDisGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'phenDis');" id="phenDis_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">Phenotype and Disease Associations</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='phenDis-4'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=gad"> GAD View
 </A><SELECT NAME="gad" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=caseControl"> Case Control
 </A><SELECT NAME="caseControl" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=nimhBipolar"> NIMH Bipolar
 </A><SELECT NAME="nimhBipolar" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=rgdQtl"> RGD Human QTL
 </A><SELECT NAME="rgdQtl" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=rgdRatQtl"> RGD Rat QTL
 </A><SELECT NAME="rgdRatQtl" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='phenDis-5'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=jaxQtlMapped"> MGI Mouse QTL
 </A><SELECT NAME="jaxQtlMapped" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="genesGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_genes_close=1#genesGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'genes');" id="genes_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">Genes and Gene Prediction Tracks</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='genes-6'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=knownGene"> UCSC Genes
 </A><SELECT NAME="knownGene" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION SELECTED>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=knownGeneOld2"> Old Known Genes
 </A><SELECT NAME="knownGeneOld2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=knownAlt"> Alt Events
 </A><SELECT NAME="knownAlt" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=ccdsGene"> CCDS
 </A><SELECT NAME="ccdsGene" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=refGene"> RefSeq Genes
 </A><SELECT NAME="refGene" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION SELECTED>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='genes-7'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=xenoRefGene"> Other RefSeq
 </A><SELECT NAME="xenoRefGene" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=mgcGenes"> MGC Genes
 </A><SELECT NAME="mgcGenes" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION SELECTED>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=orfeomeGenes"> ORFeome Clones
 </A><SELECT NAME="orfeomeGenes" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=transMap"> TransMap...
 </A><SELECT NAME="transMap" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=vegaGeneComposite"> Vega Genes
 </A><SELECT NAME="vegaGeneComposite" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='genes-8'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=ensGene"> Ensembl Genes
 </A><SELECT NAME="ensGene" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=acembly"> AceView Genes
 </A><SELECT NAME="acembly" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=sibGene"> SIB Genes
 </A><SELECT NAME="sibGene" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=nscan"> N-SCAN
 </A><SELECT NAME="nscan" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=contrastGene"> CONTRAST
 </A><SELECT NAME="contrastGene" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='genes-9'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=sgpGene"> SGP Genes
 </A><SELECT NAME="sgpGene" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=geneid"> Geneid Genes
 </A><SELECT NAME="geneid" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=genscan"> Genscan Genes
 </A><SELECT NAME="genscan" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=exoniphy"> Exoniphy
 </A><SELECT NAME="exoniphy" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=augustus"> Augustus
 </A><SELECT NAME="augustus" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='genes-10'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=rnaGene"> RNA Genes
 </A><SELECT NAME="rnaGene" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=acescan"> ACEScan
 </A><SELECT NAME="acescan" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=evofold"> EvoFold
 </A><SELECT NAME="evofold" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=wgRna"> sno/miRNA
 </A><SELECT NAME="wgRna" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=mammalPsg"> Pos Sel Genes
 </A><SELECT NAME="mammalPsg" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="rnaGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_rna_close=1#rnaGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'rna');" id="rna_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">mRNA and EST Tracks</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='rna-11'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=mrna"> Human mRNAs
 </A><SELECT NAME="mrna" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION SELECTED>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=intronEst"> Spliced ESTs
 </A><SELECT NAME="intronEst" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION SELECTED>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=est"> Human ESTs
 </A><SELECT NAME="est" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=xenoMrna"> Other mRNAs
 </A><SELECT NAME="xenoMrna" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=xenoEst"> Other ESTs
 </A><SELECT NAME="xenoEst" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='rna-12'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=HInvGeneMrna"> H-Inv
 </A><SELECT NAME="HInvGeneMrna" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=uniGene_3"> UniGene
 </A><SELECT NAME="uniGene_3" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=rnaCluster"> Gene Bounds
 </A><SELECT NAME="rnaCluster" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=sibTxGraph"> SIB Alt-Splicing
 </A><SELECT NAME="sibTxGraph" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=polyA"> Poly(A)
 </A><SELECT NAME="polyA" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='rna-13'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=cgapSage"> CGAP SAGE
 </A><SELECT NAME="cgapSage" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="regulationGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_regulation_close=1#regulationGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'regulation');" id="regulation_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">Expression and Regulation</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='regulation-14'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=affyHumanExon"> Affy All Exon
 </A><SELECT NAME="affyHumanExon" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=affyHuEx1"> Affy HuEx 1.0
 </A><SELECT NAME="affyHuEx1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=agilentCgh"> Agilent CGH
 </A><SELECT NAME="agilentCgh" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=allenBrainAli"> Allen Brain
 </A><SELECT NAME="allenBrainAli" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=gnfAtlas2"> GNF Atlas 2
 </A><SELECT NAME="gnfAtlas2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='regulation-15'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=affyRatio"> GNF Ratio
 </A><SELECT NAME="affyRatio" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=yaleBertoneTars"> Bertone Yale TAR
 </A><SELECT NAME="yaleBertoneTars" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=affyU133"> Affy U133
 </A><SELECT NAME="affyU133" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=affyGnf1h"> Affy GNF1H
 </A><SELECT NAME="affyGnf1h" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=affyU133Plus2"> Affy U133Plus2
 </A><SELECT NAME="affyU133Plus2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='regulation-16'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=affyU95"> Affy U95
 </A><SELECT NAME="affyU95" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=cpgIslandExt"> CpG Islands
 </A><SELECT NAME="cpgIslandExt" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=firstEF"> FirstEF
 </A><SELECT NAME="firstEF" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=switchDbTss"> SwitchGear TSS
 </A><SELECT NAME="switchDbTss" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=eponine"> Eponine TSS
 </A><SELECT NAME="eponine" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='regulation-17'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=vistaEnhancers"> Vista Enhancers
 </A><SELECT NAME="vistaEnhancers" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=tfbsConsSites"> TFBS Conserved
 </A><SELECT NAME="tfbsConsSites" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=wgEncodeGisSuper"> GIS PET...
 </A><SELECT NAME="wgEncodeGisSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=wgEncodeUcsdNgTaf1Super"> LI/UCSD TAF1...
 </A><SELECT NAME="wgEncodeUcsdNgTaf1Super" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=oreganno"> ORegAnno
 </A><SELECT NAME="oreganno" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='regulation-18'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=affyTxnPhase3Super"> Affy Txn...
 </A><SELECT NAME="affyTxnPhase3Super" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=regPotential7X"> 7X Reg Potential
 </A><SELECT NAME="regPotential7X" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=targetScanS"> TS miRNA sites
 </A><SELECT NAME="targetScanS" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=uppsalaChipSuper"> Uppsala ChIP...
 </A><SELECT NAME="uppsalaChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=laminB1Super"> NKI Nuc Lamina...
 </A><SELECT NAME="laminB1Super" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="compGenoGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_compGeno_close=1#compGenoGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'compGeno');" id="compGeno_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">Comparative Genomics</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='compGeno-19'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=multiz28way"> Conservation
 </A><SELECT NAME="multiz28way" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION SELECTED>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=mostConserved28way"> Most Conserved
 </A><SELECT NAME="mostConserved28way" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=multiz17way"> 17-Way Cons
 </A><SELECT NAME="multiz17way" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=phastConsElements17way"> 17-Way Most Cons
 </A><SELECT NAME="phastConsElements17way" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=consIndelsHgMmCanFam"> Cons Indels MmCf
 </A><SELECT NAME="consIndelsHgMmCanFam" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-20'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainStrPur2"> S. purpuratus Chain
 </A><SELECT NAME="chainStrPur2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netStrPur2"> S. purpuratus Net
 </A><SELECT NAME="netStrPur2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainBraFlo1"> Lancelet Chain
 </A><SELECT NAME="chainBraFlo1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netBraFlo1"> Lancelet Net
 </A><SELECT NAME="netBraFlo1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainPetMar1"> Lamprey Chain
 </A><SELECT NAME="chainPetMar1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-21'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netPetMar1"> Lamprey Net
 </A><SELECT NAME="netPetMar1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainOryLat1"> Medaka Chain
 </A><SELECT NAME="chainOryLat1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netOryLat1"> Medaka Net
 </A><SELECT NAME="netOryLat1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainGasAcu1"> Stickleback Chain
 </A><SELECT NAME="chainGasAcu1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netGasAcu1"> Stickleback Net
 </A><SELECT NAME="netGasAcu1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-22'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainFr2"> Fugu Chain
 </A><SELECT NAME="chainFr2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netFr2"> Fugu Net
 </A><SELECT NAME="netFr2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainTetNig1"> Tetraodon Chain
 </A><SELECT NAME="chainTetNig1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netTetNig1"> Tetraodon Net
 </A><SELECT NAME="netTetNig1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=ecoresTetNig1"> Tetraodon Ecores
 </A><SELECT NAME="ecoresTetNig1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-23'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainDanRer5"> Zebrafish Chain
 </A><SELECT NAME="chainDanRer5" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netDanRer5"> Zebrafish Net
 </A><SELECT NAME="netDanRer5" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainXenTro2"> X. tropicalis Chain
 </A><SELECT NAME="chainXenTro2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netXenTro2"> X. tropicalis Net
 </A><SELECT NAME="netXenTro2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainGalGal3"> Chicken Chain
 </A><SELECT NAME="chainGalGal3" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-24'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netGalGal3"> Chicken Net
 </A><SELECT NAME="netGalGal3" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainAnoCar1"> Lizard Chain
 </A><SELECT NAME="chainAnoCar1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netAnoCar1"> Lizard Net
 </A><SELECT NAME="netAnoCar1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainOrnAna1"> Platypus Chain
 </A><SELECT NAME="chainOrnAna1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netOrnAna1"> Platypus Net
 </A><SELECT NAME="netOrnAna1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-25'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainMonDom4"> Opossum Chain
 </A><SELECT NAME="chainMonDom4" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netMonDom4"> Opossum Net
 </A><SELECT NAME="netMonDom4" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainCanFam2"> Dog Chain
 </A><SELECT NAME="chainCanFam2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netCanFam2"> Dog Net
 </A><SELECT NAME="netCanFam2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainFelCat3"> Cat Chain
 </A><SELECT NAME="chainFelCat3" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-26'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netFelCat3"> Cat Net
 </A><SELECT NAME="netFelCat3" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainEquCab1"> equCab1 Chain
 </A><SELECT NAME="chainEquCab1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netEquCab1"> equCab1 Net
 </A><SELECT NAME="netEquCab1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainBosTau4"> Cow Chain
 </A><SELECT NAME="chainBosTau4" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netBosTau4"> Cow Net
 </A><SELECT NAME="netBosTau4" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-27'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainCavPor3"> Guinea Pig Chain
 </A><SELECT NAME="chainCavPor3" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netCavPor3"> Guinea Pig Net
 </A><SELECT NAME="netCavPor3" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainRn4"> Rat Chain
 </A><SELECT NAME="chainRn4" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netRn4"> Rat Net
 </A><SELECT NAME="netRn4" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainMm9"> Mouse Chain
 </A><SELECT NAME="chainMm9" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-28'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netMm9"> Mouse Net
 </A><SELECT NAME="netMm9" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainCalJac1"> Marmoset Chain
 </A><SELECT NAME="chainCalJac1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netCalJac1"> Marmoset Net
 </A><SELECT NAME="netCalJac1" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainRheMac2"> Rhesus Chain
 </A><SELECT NAME="chainRheMac2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netRheMac2"> Rhesus Net
 </A><SELECT NAME="netRheMac2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='compGeno-29'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainPonAbe2"> Orangutan Chain
 </A><SELECT NAME="chainPonAbe2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netPonAbe2"> Orangutan Net
 </A><SELECT NAME="netPonAbe2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainPanTro2"> Chimp Chain
 </A><SELECT NAME="chainPanTro2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=netPanTro2"> Chimp Net
 </A><SELECT NAME="netPanTro2" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="varRepGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_varRep_close=1#varRepGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'varRep');" id="varRep_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">Variation and Repeats</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='varRep-30'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=snp129"> SNPs (129)
 </A><SELECT NAME="snp129" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION SELECTED>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=snp128"> SNPs (128)
 </A><SELECT NAME="snp128" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION SELECTED>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=snp126"> SNPs (126)
 </A><SELECT NAME="snp126" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=snpArray"> SNP Arrays
 </A><SELECT NAME="snpArray" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=hapmapSnps"> HapMap SNPs
 </A><SELECT NAME="hapmapSnps" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='varRep-31'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=hapmapLdPh"> HapMap LD Phased
 </A>[No data-chrX]</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=tajdSnp"> Tajima's D SNPs
 </A><SELECT NAME="tajdSnp" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=tajD"> Tajima's D
 </A><SELECT NAME="tajD" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=dgv"> DGV Struct Var
 </A><SELECT NAME="dgv" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=genomicSuperDups"> Segmental Dups
 </A><SELECT NAME="genomicSuperDups" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='varRep-32'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=cnp"> Structural Var
 </A><SELECT NAME="cnp" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=exaptedRepeats"> Exapted Repeats
 </A><SELECT NAME="exaptedRepeats" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=rmsk"> RepeatMasker
 </A><SELECT NAME="rmsk" class=normalText style="width: 70px" >

<OPTION>hide</OPTION>

<OPTION SELECTED>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=nestedRepeats"> Interrupted Rpts
 </A><SELECT NAME="nestedRepeats" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=simpleRepeat"> Simple Repeats
 </A><SELECT NAME="simpleRepeat" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='varRep-33'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=microsat"> Microsatellite
 </A><SELECT NAME="microsat" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=chainSelf"> Self Chain
 </A><SELECT NAME="chainSelf" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="encodeGenesGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_encodeGenes_close=1#encodeGenesGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'encodeGenes');" id="encodeGenes_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">ENCODE Regions and Genes</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='encodeGenes-34'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeRegions"> ENCODE Regions
 </A><SELECT NAME="encodeRegions" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeGencodeSuper"> Gencode Genes...
 </A><SELECT NAME="encodeGencodeSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeEgaspSuper"> EGASP...
 </A><SELECT NAME="encodeEgaspSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodePseudogene"> Pseudogenes
 </A><SELECT NAME="encodePseudogene" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeRna"> Known+Pred RNA
 </A><SELECT NAME="encodeRna" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='encodeGenes-35'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeUViennaRnaz"> Vienna RNAz
 </A><SELECT NAME="encodeUViennaRnaz" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="encodeTxLevelsGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_encodeTxLevels_close=1#encodeTxLevelsGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'encodeTxLevels');" id="encodeTxLevels_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">ENCODE Transcript Levels</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='encodeTxLevels-36'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeAffyRnaSuper"> Affy RNA...
 </A><SELECT NAME="encodeAffyRnaSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeBuFirstExon"> BU First Exon
 </A><SELECT NAME="encodeBuFirstExon" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeRikenCage"> Riken CAGE
 </A><SELECT NAME="encodeRikenCage" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeStanfordPromoters"> Stanf Promoter
 </A><SELECT NAME="encodeStanfordPromoters" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeStanfordRtPcr"> Stanf RTPCR
 </A><SELECT NAME="encodeStanfordRtPcr" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='encodeTxLevels-37'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeYaleRnaSuper"> Yale RNA...
 </A><SELECT NAME="encodeYaleRnaSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeAffyEcSuper"> Affy EC...
 </A>[No data-chrX]</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="encodeChipGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_encodeChip_close=1#encodeChipGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'encodeChip');" id="encodeChip_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">ENCODE Chromatin Immunoprecipitation</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='encodeChip-38'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeAffyChipSuper"> Affy ChIP...
 </A><SELECT NAME="encodeAffyChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeUcsdChipSuper"> LI/UCSD ChIP...
 </A><SELECT NAME="encodeUcsdChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeSangerChipSuper"> Sanger ChIP-chip...
 </A><SELECT NAME="encodeSangerChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeStanfordChipSuper"> Stanf ChIP...
 </A><SELECT NAME="encodeStanfordChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeUcDavisChipSuper"> UC Davis ChIP...
 </A><SELECT NAME="encodeUcDavisChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td></tr><tr style='display: ' id='encodeChip-39'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeUtexChipSuper"> UT-Austin ChIP...
 </A><SELECT NAME="encodeUtexChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeUppsalaChipSuper"> Uppsala ChIP...
 </A><SELECT NAME="encodeUppsalaChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeYaleChipSuper"> Yale ChIP...
 </A><SELECT NAME="encodeYaleChipSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="encodeChromGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_encodeChrom_close=1#encodeChromGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'encodeChrom');" id="encodeChrom_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">ENCODE Chromosome, Chromatin and DNA Structure</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='encodeChrom-40'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeBUORChID"> BU ORChID
 </A><SELECT NAME="encodeBUORChID" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeNhgriDukeDnaseHs"> Duke/NHGRI DNase
 </A><SELECT NAME="encodeNhgriDukeDnaseHs" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeUncFaire"> UNC FAIRE
 </A><SELECT NAME="encodeUncFaire" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeUvaDnaRepSuper"> UVa DNA Rep...
 </A><SELECT NAME="encodeUvaDnaRepSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeUwDnaseSuper"> UW DNase...
 </A><SELECT NAME="encodeUwDnaseSuper" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>show</OPTION>

</SELECT>

</td></tr><TR><th align="left" colspan=5 BGCOLOR=#536ED3><table width='100%'><tr><td align='left'>

<A NAME="encodeCompAndVarGroup"></A><A HREF="../cgi-bin/hgTracks?hgsid=109992416&hgtgroup_encodeCompAndVar_close=1#encodeCompAndVarGroup" class='bigBlue'><IMG height='18' width='18' onclick="return toggleTrackGroupVisibility(this, 'encodeCompAndVar');" id="encodeCompAndVar_button" src="../images/remove_sm.gif" alt="-" class='bigBlue'></A> </td><td align='center' width='100%'>

<B><FONT COLOR="#FFFFFF">ENCODE Comparative Genomics and Variation</FONT></B></td><td align='right'>

<input type='submit' name='submit' value='refresh'>

</td></tr></table></th>

</tr><tr style='display: ' id='encodeCompAndVar-41'><td align=left><A HREF="../cgi-bin/hgTrackUi?hgsid=109992416&c=chrX&g=encodeIndels"> NHGRI DIPs
 </A><SELECT NAME="encodeIndels" class=hiddenText style="width: 70px" >

<OPTION SELECTED>hide</OPTION>

<OPTION>dense</OPTION>

<OPTION>squish</OPTION>

<OPTION>pack</OPTION>

<OPTION>full</OPTION>

</SELECT>

</td></tr></table>

</DIV>

<INPUT TYPE=SUBMIT NAME="submit" VALUE="refresh" ></CENTER>

</FORM>

<FORM ACTION='../cgi-bin/hgCustom' NAME='customTrackForm'><INPUT TYPE=HIDDEN NAME="hgsid" VALUE="109992416"></FORM>

</BODY>

</HTML>

Appendix 32 – UCSC Track for chromosome 9

browser position chr9:27357278-27579657

track name=Fin_Hap description="ALS Finland HAP" color=0,0,255,

chr9	27357278	27357278	rs1330921
chr9	27375002	27375002	rs1110264
chr9	27376505	27376505	rs1110155
chr9	27392961	27392961	rs2150336
chr9	27399264	27399264	rs2225389
chr9	27420232	27420232	rs1161680
chr9	27428411	27428411	rs2120718
chr9	27433802	27433802	rs2589054
chr9	27435104	27435104	rs10812596
chr9	27437089	27437089	rs1058326
chr9	27440211	27440211	rs944404
chr9	27442240	27442240	rs1316679
chr9	27448939	27448939	rs725804
chr9	27458461	27458461	rs10511816
chr9	27467874	27467874	rs1444533
chr9	27468052	27468052	rs1822723
chr9	27472235	27472235	rs4879515
chr9	27473959	27473959	rs895023
chr9	27480967	27480967	rs7046653
chr9	27485418	27485418	rs2440622
chr9	27492986	27492986	rs1977661
chr9	27519316	27519316	rs903603
chr9	27526397	27526397	rs2814707
chr9	27533281	27533281	rs3849942
chr9	27546780	27546780	rs10122902
chr9	27547919	27547919	rs10757665
chr9	27551049	27551049	rs774359
chr9	27562255	27562255	rs2282241
chr9	27565785	27565785	rs1948522
chr9	27569560	27569560	rs1982915

chr9 27578731 27578731 rs702231

chr9 27579657 27579657 rs696826

track name=Common_Hap description="ALS UK_FRA_HOL_US HAP" color=255,0,0,

chr9 27357278 27357278 rs1330921

chr9 27375002 27375002 rs1110264

chr9 27376505 27376505 rs1110155

chr9 27392961 27392961 rs2150336

chr9 27399264 27399264 rs2225389

chr9 27420232 27420232 rs1161680

chr9 27428411 27428411 rs2120718

chr9 27433802 27433802 rs2589054

chr9 27435104 27435104 rs10812596

chr9 27437089 27437089 rs1058326

chr9 27440211 27440211 rs944404

chr9 27442240 27442240 rs1316679

chr9 27448939 27448939 rs725804

chr9 27458461 27458461 rs10511816

chr9 27467874 27467874 rs1444533

chr9 27468052 27468052 rs1822723

chr9 27472235 27472235 rs4879515

chr9 27473959 27473959 rs895023

chr9 27480967 27480967 rs7046653

chr9 27485418 27485418 rs2440622

chr9 27492986 27492986 rs1977661

chr9 27519316 27519316 rs903603

chr9 27526397 27526397 rs2814707

chr9 27533281 27533281 rs3849942

chr9 27546780 27546780 rs10122902

chr9 27547919 27547919 rs10757665

chr9 27551049 27551049 rs774359

chr9 27562255 27562255 rs2282241

chr9 27565785 27565785 rs1948522

chr9 27569560 27569560 rs1982915

chr9 27578731 27578731 rs702231

chr9 27579657 27579657 rs696826

track name=BE_Hap description="ALS BE HAP" color=55,0,0,

chr9 27357278 27357278 rs1330921

chr9 27375002 27375002 rs1110264

chr9 27376505 27376505 rs1110155

chr9 27392961 27392961 rs2150336

chr9 27399264 27399264 rs2225389

chr9 27420232 27420232 rs1161680

chr9 27428411 27428411 rs2120718

chr9 27433802 27433802 rs2589054

chr9 27435104 27435104 rs10812596

chr9 27437089 27437089 rs1058326

chr9 27440211 27440211 rs944404

chr9 27442240 27442240 rs1316679

chr9 27448939 27448939 rs725804

track name=SWE_Hap description="ALS SWE HAP" color=55,0,0,

chr9 27458461 27458461 rs10511816

chr9 27467874 27467874 rs1444533

chr9 27468052 27468052 rs1822723

chr9 27472235 27472235 rs4879515

chr9 27473959 27473959 rs895023

chr9 27480967 27480967 rs7046653

chr9 27485418 27485418 rs2440622

chr9 27492986 27492986 rs1977661

chr9 27519316 27519316 rs903603

chr9 27526397 27526397 rs2814707

chr9 27533281 27533281 rs3849942

chr9 27546780 27546780 rs10122902

chr9 27547919 27547919 rs10757665

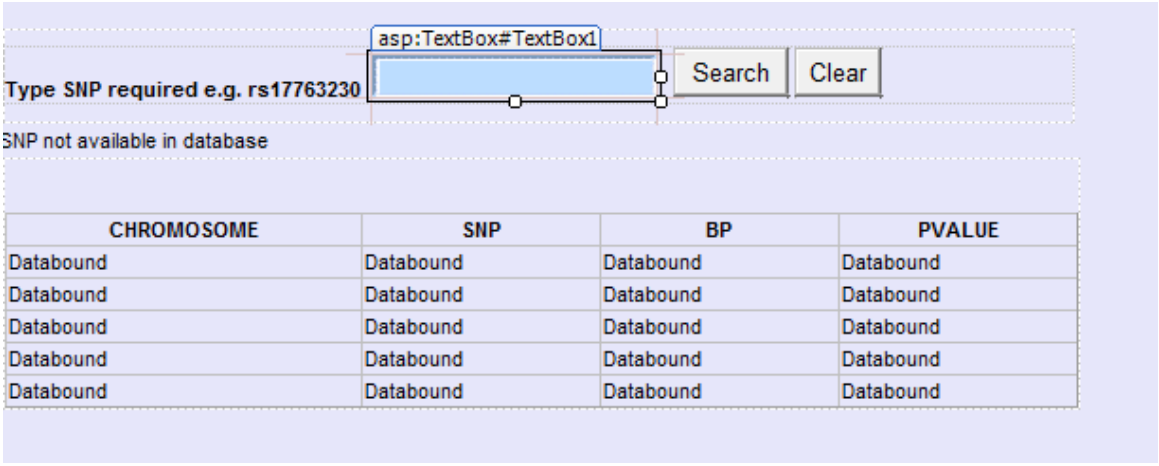
chr9 27551049 27551049 rs774359

chr9 27562255 27562255 rs2282241

chr9 27565785 27565785 rs1948522

chr9	27569560	27569560	rs1982915
chr9	27578731	27578731	rs702231
chr9	27579657	27579657	rs696826

Appendix 33 – Codes for searching SNP ids on GWAS page




```
Protected Sub Search_Click(ByVal sender As Object, ByVal e As
System.EventArgs)
        Label1.Visible= "False"
        If TextBox1.Text = "" Then
          Response.Redirect("~/GWA2/gwas_fogh.aspx")
        ElseIf TextBox1.Text <> "" Then
            Dim dv1 As DataView
            dv1 =
CType(SqlDataSource1.Select(DataSourceSelectArguments.Empty), DataView)
            If (dv1.Table.Rows.Count = 0) Then
            Label1.Visible= "True"
            ElseIf      (dv1.Table.Rows.Count > 0) Then
            Dim SNP As String = TextBox1.Text
            GridView1.Visible = "True"
            End If

        End If
    End Sub
```

Where:

```
SqlDataSource1 = SELECT DISTINCT [SNP] FROM dbo.[GWA_FOGH] WHERE ([SNP] =
@SNP)
```

Appendix 34 – Credibility analysis of genes on SPSS and spreadsheet

Scale: Reliability Analysis using Cronbach's alpha

ALSoD_credibility.sav [DataSet1] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : Gene | SOD1

	Gene	ALSoD	Forced	Unforced	var
1	SOD1	1.00	1.00	1.00	
2	TARDBP(TDP43)	2.00	2.00	1.00	
3	ANG	3.00	5.00	9.00	
4	FUS	4.00	3.00	1.00	
5	OPTN	5.00	4.00	5.00	
6	ALS2	6.00	12.00	8.00	
7	NEFH	7.00	8.00	13.00	
8	SETX	8.00	10.00	6.00	
9	FIG4	9.00	11.00	10.00	
10	VCP	10.00	6.00	4.00	
11	DCTN1	11.00	13.00	12.00	
12	VAPB	12.00	7.00	6.00	
13	DAO	13.00	9.00	11.00	
14					

Data View | Variable View

SPSS Processor is ready

Case Processing Summary

		N	%
Cases	Valid	13	100.0
	Excluded[a]	0	.0
	Total	13	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.865	.866	3

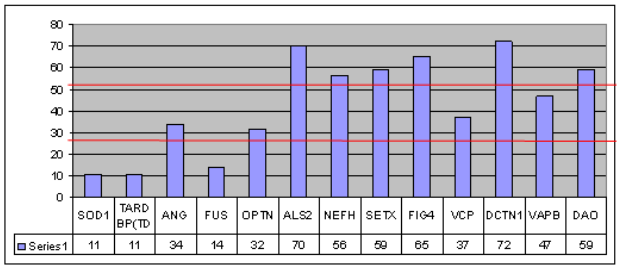
Inter-Item Correlation Matrix

	ALSoD	Forced	Unforced
ALSoD	1.000	.692	.579
Forced	.692	1.000	.778
Unforced	.579	.778	1.000

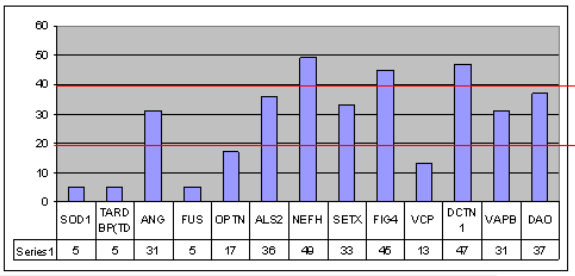
Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
ALSoD	13.6923	58.397	.672	.484	.873
Forced	13.6923	51.897	.829	.693	.732
Unforced	14.0000	51.333	.738	.608	.818

A high Cronbach's Alpha of 0.8 shows that there's a high reliability and correlation between the three methods of testing credibility.

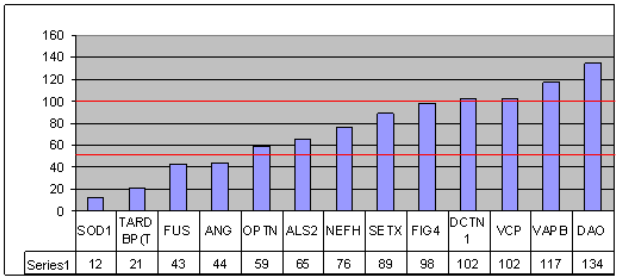


Forced



Non Forced

The lower the value, the more believable the gene.



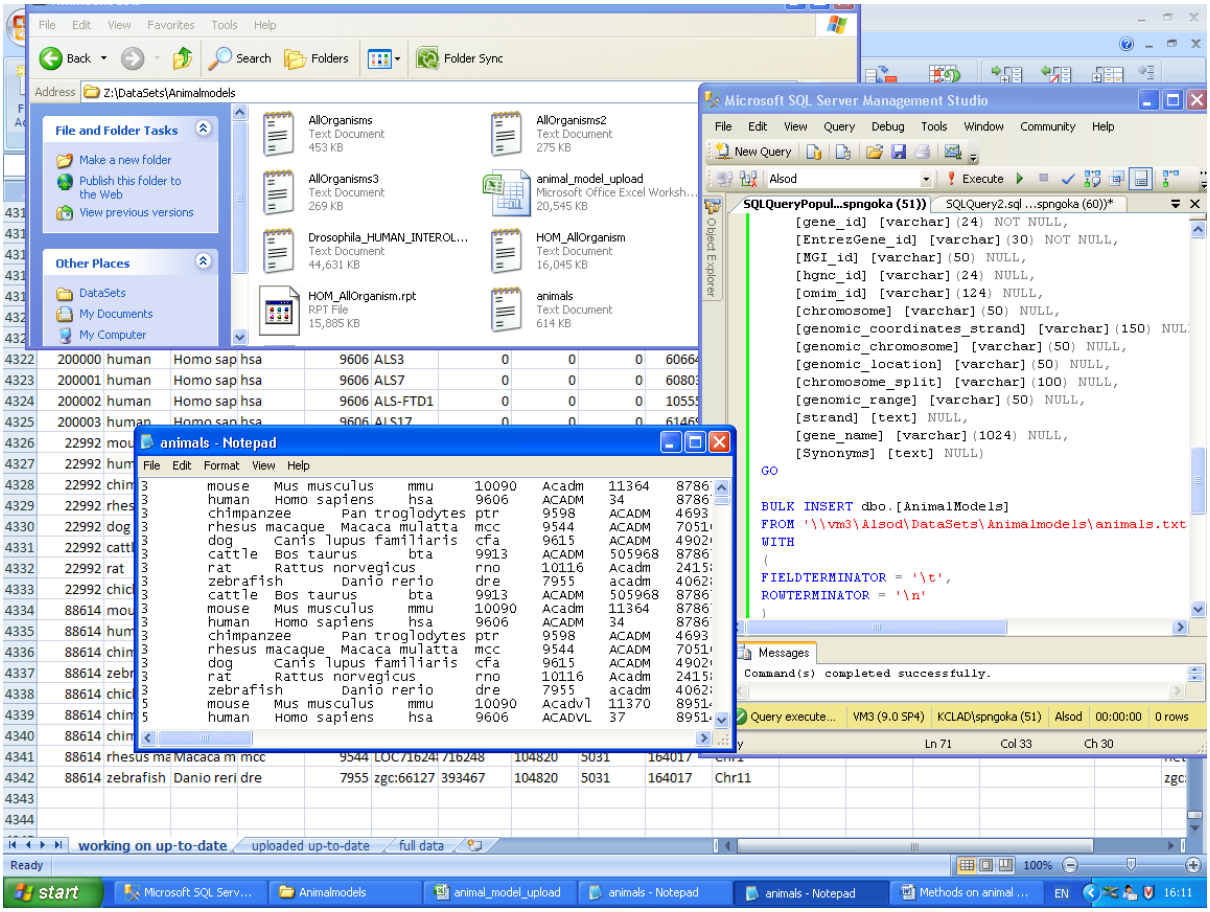
ALSoD

First Five Credible Genes

- SOD1
- TARDBP (TDP43)
- FUS
- OPTN
- ANG

Appendix 35 – Populating Animal model table using T-SQL

```
CREATE TABLE [dbo].[AnimalModels](
	[HomoloGeneID] [varchar](50) NULL,
	[Organism_Name] [varchar](30) NOT NULL,
	[Biological_Name] [varchar](100) NOT NULL,
	[Kegg_id] [varchar](30) NULL,
	[ncbi_locuslink_id] [varchar](50) NULL,
	[gene_id] [varchar](24) NOT NULL,
	[EntrezGene_id] [varchar](30) NOT NULL,
	[MGI_id] [varchar](50) NULL,
	[hgnc_id] [varchar](24) NULL,
	[omim_id] [varchar](124) NULL,
	[chromosome] [varchar](50) NULL,
	[genomic_coordinates_strand] [varchar](150) NULL,
	[genomic_chromosome] [varchar](50) NULL,
	[genomic_location] [varchar](50) NULL,
	[chromosome_split] [varchar](100) NULL,
	[genomic_range] [varchar](50) NULL,
	[strand] [text] NULL,
	[gene_name] [varchar](1024) NULL,
	[Synonyms] [text] NULL)
GO

BULK INSERT dbo.[AnimalModels]
FROM '\\vm3\Alsod\DataSets\Animalmodels\animals.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.[AnimalModels]
ADD [ID] [int] IDENTITY(1,1) NOT NULL;
GO
```



```
--Importing reference data from MGI
CREATE TABLE dbo.[Reference_MGI](
      [MGI_id] [varchar](50) NULL,
      [gene_id] [varchar](64) NOT NULL,
      [pubmed_id] [varchar](20) NOT NULL)
GO

BULK INSERT dbo.[Reference_MGI]
FROM '\\vm3\Alsod\DataSets\Animalmodels\MRK_Reference_uploadversion.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.[Reference_MGI]
ADD [ID] [int] IDENTITY(1,1) PRIMARY KEY;
GO

--Importing reference data from pubmed
CREATE TABLE dbo.[Reference_Pubmed](
      [title] [varchar](1024) NULL,
      [pubmed_id] [varchar](20) NOT NULL,
      [year] [varchar](5) NOT NULL,
      [first_author] [varchar](24) NOT NULL)
GO

BULK INSERT dbo.[Reference_Pubmed]
```

```
FROM '\\vm3\Alsod\DataSets\Animalmodels\pubmed_references_oct13.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.[Reference_Pubmed]
ADD [ID] [int] IDENTITY(1,1) PRIMARY KEY;
GO

CREATE TABLE dbo.[Animal_Pubmed](
      [title] [varchar](1024) NOT NULL,
      [pubmed_id] [varchar](20) NOT NULL,
      [year] [varchar](5) NOT NULL,
      [first_author] [varchar](24) NOT NULL,
      [EntrezGene_id] [varchar](30) NOT NULL,
      [gene_id] [varchar](64) NOT NULL)
GO

BULK INSERT dbo.[Animal_Pubmed]
FROM '\\vm3\Alsod\DataSets\Animalmodels\animal_spast.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

--After importing all, then:
ALTER TABLE dbo.[Animal_Pubmed]
ADD [ID] [int] IDENTITY(1,1) PRIMARY KEY;
GO
```

Appendix 36 – GWAS method for 5 Populations

```
USE [Alsod]
GO

---Data for UK
CREATE TABLE dbo.[GWA_UK](
	[CHROMOSOME] [int] NULL,
	[SNP] [varchar](30) PRIMARY KEY,
	[BP] [int] NULL,
	[A1] [varchar](1) NULL,
	[F_A] [float] NULL,
	[F_U] [float] NULL,
	[A2] [varchar](1) NULL,
	[CHISQ] [float] NULL,
	[PVALUE] [float] NULL,
	[ODDRATIO] [float] NULL,
	[L95] [float] NULL,
	[U95] [float] NULL)
GO

BULK INSERT dbo.[GWA_UK]
FROM '\\vm3\Alsod\DataSets\uk\uk_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO

---Data for BOSTON
CREATE TABLE dbo.[GWA_BOS](
	[CHROMOSOME] [int] NULL,
	[SNP] [varchar](30) PRIMARY KEY,
	[BP] [int] NULL,
	[A1] [varchar](1) NULL,
	[F_A] [float] NULL,
	[F_U] [float] NULL,
	[A2] [varchar](1) NULL,
	[CHISQ] [float] NULL,
	[PVALUE] [float] NULL,
	[ODDRATIO] [float] NULL,
	[L95] [float] NULL,
	[U95] [float] NULL)
GO

BULK INSERT dbo.[GWA_BOS]
FROM '\\vm3\Alsod\DataSets\bos\bos_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO

--Data for HOLLAND
CREATE TABLE dbo.[GWA_HOL](
	[CHROMOSOME] [int] NULL,
	[SNP] [varchar](30) PRIMARY KEY,
	[BP] [int] NULL,
	[A1] [varchar](1) NULL,
	[F_A] [float] NULL,
```

```
	[F_U] [float] NULL,
	[A2] [varchar](1) NULL,
	[CHISQ] [float] NULL,
	[PVALUE] [float] NULL,
	[ODDRATIO] [float] NULL,
	[L95] [float] NULL,
	[U95] [float] NULL)
GO
BULK INSERT dbo.[GWA_HOL]
FROM '\\vm3\Alsod\DataSets\hol\hol_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO

---Data for FRANCE
CREATE TABLE dbo.[GWA_FRA](
	[CHROMOSOME] [int] NULL,
	[SNP] [varchar](30) PRIMARY KEY,
	[BP] [int] NULL,
	[A1] [varchar](1) NULL,
	[F_A] [float] NULL,
	[F_U] [float] NULL,
	[A2] [varchar](1) NULL,
	[CHISQ] [float] NULL,
	[PVALUE] [float] NULL,
	[ODDRATIO] [float] NULL,
	[L95] [float] NULL,
	[U95] [float] NULL)
GO
BULK INSERT dbo.[GWA_FRA]
FROM '\\vm3\Alsod\DataSets\fra\fra_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO

--Data for USA
CREATE TABLE dbo.[GWA_USA](
	[CHROMOSOME] [int] NULL,
	[SNP] [varchar](30) PRIMARY KEY,
	[BP] [int] NULL,
	[A1] [varchar](1) NULL,
	[F_A] [float] NULL,
	[F_U] [float] NULL,
	[A2] [varchar](1) NULL,
	[CHISQ] [float] NULL,
	[PVALUE] [float] NULL,
	[ODDRATIO] [float] NULL,
	[L95] [float] NULL,
	[U95] [float] NULL)
GO
BULK INSERT dbo.[GWA_USA]
FROM '\\vm3\Alsod\DataSets\usa\usa_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
```

```
)
GO

--------neglogpvalue for single tables
ALTER TABLE dbo.GWA_UK
ADD neglogpval AS (-1 * log10(PVALUE)) PERSISTED;
GO

ALTER TABLE dbo.GWA_BOS
ADD neglogpval AS (-1 * log10(PVALUE)) PERSISTED;
GO

ALTER TABLE dbo.GWA_HOL
ADD neglogpval AS (-1 * log10(PVALUE)) PERSISTED;
GO

ALTER TABLE dbo.GWA_FRA
ADD neglogpval AS (-1 * log10(PVALUE)) PERSISTED;
GO

ALTER TABLE dbo.GWA_USA
ADD neglogpval AS (-1 * log10(PVALUE)) PERSISTED;
GO

--Create table to insert all data
CREATE TABLE dbo.[GWA_NOKEY]
([CHROMOSOME] [int] NULL,
      [SNP] [varchar](30) NULL,
      [BP] [int] NULL,
      [A1] [varchar](1) NULL,
      [F_A] [float] NULL,
      [F_U] [float] NULL,
      [A2] [varchar](1) NULL,
      [CHISQ] [float] NULL,
      [PVALUE] [float] NULL,
      [ODDRATIO] [float] NULL,
      [L95] [float] NULL,
      [U95] [float] NULL)
GO

---Bulk Insert the data from space delimited file
BULK INSERT dbo.[GWA_NOKEY]
FROM '\\vm3\Alsod\DataSets\uk\uk_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO


BULK INSERT dbo.GWA_NOKEY
FROM '\\vm3\Alsod\DataSets\bos\bos_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO

BULK INSERT dbo.GWA_NOKEY
FROM '\\vm3\Alsod\DataSets\hol\hol_result_assoc.txt'
```

```
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO

BULK INSERT dbo.GWA_NOKEY
FROM '\\vm3\Alsod\DataSets\fra\fra_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO

BULK INSERT dbo.GWA_NOKEY
FROM '\\vm3\Alsod\DataSets\usa\usa_result_assoc.txt'
WITH
(
FIELDTERMINATOR = ' ',
ROWTERMINATOR = '\n'
)
GO

---Insert ID and Population column
ALTER TABLE dbo.GWA_NOKEY
ADD ID INT IDENTITY(1,1);
GO

ALTER TABLE dbo.GWA_NOKEY
ADD POP VARCHAR(50);
GO

UPDATE dbo.GWA_NOKEY SET POP = 'UK'
WHERE ID BETWEEN 1 AND 275619
GO

UPDATE dbo.GWA_NOKEY SET POP = 'BOSTON'
WHERE ID BETWEEN 275620 AND 558067
GO

UPDATE dbo.GWA_NOKEY SET POP = 'HOLLAND'
WHERE ID BETWEEN 558068 AND 846203
GO

UPDATE dbo.GWA_NOKEY SET POP = 'FRANCE'
WHERE ID BETWEEN 846204 AND 1132710
GO

UPDATE dbo.GWA_NOKEY SET POP = 'USA'
WHERE ID BETWEEN 1132711 AND 1420350
GO

-----------From Haploview format to database

USE [Alsod]
GO

---UK, BOSTON
create table dbo.GWA_UK_BOS
(chr int NOT NULL,
```

```
snp varchar(30) NOT NULL,
bp int NOT NULL,
pvalbos float NULL,
pvaluk float NULL,
newpvalueukbos float NULL)
go


BULK INSERT dbo.GWA_UK_BOS
FROM '\\vm3\Alsod\DataSets\DataPlots\uk_bos.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.GWA_UK_BOS
ADD id INT IDENTITY(1,1) PRIMARY KEY;
GO

------UK,BOSTON,FRANCE
create table dbo.GWA_UK_BOS_FRA
(chr int NOT NULL,
snp varchar(30) NOT NULL,
bp int NOT NULL,
pvalfra float NULL,
pvalbos float NULL,
newpvaluebosfra float NULL,
pvaluk float NULL,
newpvalueukbosfra float NULL)
go


BULK INSERT dbo.GWA_UK_BOS_FRA
FROM '\\vm3\Alsod\DataSets\DataPlots\uk_bos_fra.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA
ADD id INT IDENTITY(1,1) PRIMARY KEY;
GO

------UK,BOSTON,FRANCE,USA
create table dbo.GWA_UK_BOS_FRA_USA
(chr int NOT NULL,
snp varchar(30) NOT NULL,
bp int NOT NULL,
pvalusa float NULL,
pvalfra float NULL,
newpvaluefrausa float NULL,
pvalbos float NULL,
newpvaluebosfrausa float NULL,
pvaluk float NULL,
newpvalueukbosfrausa float NULL,
newpvalueukfrausa float NULL,
newpvalueukfra float NULL,
newpvaluebosusa float NULL,
```

```
newpvalueukusa float NULL)
go


BULK INSERT dbo.GWA_UK_BOS_FRA_USA
FROM '\\vm3\Alsod\DataSets\DataPlots\uk_bos_fra_usa.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA_USA
ADD id INT IDENTITY(1,1) PRIMARY KEY;
GO

------UK,BOSTON,HOLLAND,FRANCE,USA
create table dbo.GWA_UK_BOS_HOL_FRA_USA
(chr int NOT NULL,
snp varchar(30) NOT NULL,
bp int NOT NULL,
pvalhol float NULL,
pvalusa float NULL,
newpvalueholusa float NULL,
pvalfra float NULL,
newpvalueholfrausa float NULL,
pvalbos float NULL,
newpvaluebosholfrausa float NULL,
pvaluk float NULL,
newpvalueukbosholfrausa float NULL,
newpvalueholfra float NULL,
newpvalueboshol float NULL,
newpvalueukhol float NULL,
newpvaluebosholusa float NULL,
newpvaluebosholfra float NULL,
newpvalueukholusa float NULL,
newpvalueukholfra float NULL,
newpvalueukbosusa float NULL,
newpvalueukboshol float NULL,
newpvalueukbosholfra float NULL,
newpvalueukbosholusa float NULL,
newpvalueukholfrausa float NULL)
go


BULK INSERT dbo.GWA_UK_BOS_HOL_FRA_USA
FROM '\\vm3\Alsod\DataSets\DataPlots\uk_bos_hol_fra_usa.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD id INT IDENTITY(1,1) PRIMARY KEY;
GO

---Add log of pvalue column to plot graph
ALTER TABLE dbo.GWA_UK_BOS
ADD log_ukbos AS (-1 * log10(newpvalueukbos)) PERSISTED;
```

```
GO
--
ALTER TABLE dbo.GWA_UK_BOS_FRA
ADD log_bosfra AS (-1 * log10(newpvaluebosfra)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA
ADD log_ukbosfra AS (-1 * log10(newpvalueukbosfra)) PERSISTED;
GO
--
ALTER TABLE dbo.GWA_UK_BOS_FRA_USA
ADD log_frausa AS (-1 * log10(newpvaluefrausa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA_USA
ADD log_bosfrausa AS (-1 * log10(newpvaluebosfrausa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA_USA
ADD log_ukbosfrausa AS (-1 * log10(newpvalueukbosfrausa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA_USA
ADD log_ukfrausa AS (-1 * log10(newpvalueukfrausa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA_USA
ADD log_ukfra AS (-1 * log10(newpvalueukfra)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA_USA
ADD log_bosusa AS (-1 * log10(newpvaluebosusa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_FRA_USA
ADD log_ukusa AS (-1 * log10(newpvalueukusa)) PERSISTED;
GO
----
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_holusa AS (-1 * log10(newpvalueholusa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_holfrausa AS (-1 * log10(newpvalueholfrausa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_bosholfrausa AS (-1 * log10(newpvaluebosholfrausa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_ukbosholfrausa AS (-1 * log10(newpvalueukbosholfrausa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_holfra AS (-1 * log10(newpvalueholfra)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_boshol AS (-1 * log10(newpvalueboshol)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
```

```
ADD log_ukhol AS (-1 * log10(newpvalueukhol)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_bosholusa AS (-1 * log10(newpvaluebosholusa)) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_bosholfra AS (-1 * log10(newpvaluebosholfra)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_ukholusa AS (-1 * log10(newpvalueukholusa)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_ukholfra AS (-1 * log10(newpvalueukholfra)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_ukbosusa AS (-1 * log10(newpvalueukbosusa)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_ukboshol AS (-1 * log10(newpvalueukboshol)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_ukbosholfra AS (-1 * log10(newpvalueukbosholfra)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_ukbosholusa AS (-1 * log10(newpvalueukbosholusa)) PERSISTED;
GO
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD log_ukholfrausa AS (-1 * log10(newpvalueukholfrausa)) PERSISTED;
GO

---negative 2 log of each p-value
ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD neg2logpvalhol AS -2 * log(pvalhol) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD neg2logpvalusa AS -2 * log(pvalusa) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD neg2logpvalfra AS -2 * log(pvalfra) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD neg2logpvalbos AS -2 * log(pvalbos) PERSISTED;
GO

ALTER TABLE dbo.GWA_UK_BOS_HOL_FRA_USA
ADD neg2logpvaluk AS -2 * log(pvaluk) PERSISTED;
GO

ALTER TABLE dbo.ANALYZE_THEIRDATA
ADD neg2logpval AS -2 * log(PVALUE) PERSISTED;
GO

---
CREATE TABLE [dbo].[ANALYZE_OURDATA](
      [CHROMOSOME] [int] NULL,
      [SNP] [varchar](30) PRIMARY KEY,
      [BP] [int] NULL,
      [PVALHOL] [float] NULL,
```

```
      [PVALUSA] [float] NULL,
      [PVALFRA] [float] NULL,
      [PVALBOS] [float] NULL,
      [PVALUK] [float] NULL)
GO


INSERT INTO dbo.ANALYZE_OURDATA (CHROMOSOME, SNP, BP, PVALHOL, PVALUSA,
PVALFRA, PVALBOS, PVALUK)
SELECT chr, snp, bp, pvalhol, pvalusa, pvalfra, pvalbos, pvaluk FROM
dbo.GWA_UK_BOS_HOL_FRA_USA
GO

--calculate -2ln(p) on ourdata
ALTER TABLE dbo.ANALYZE_OURDATA
ADD neg2logpvalhol AS -2 * log(pvalhol) PERSISTED;
GO

ALTER TABLE dbo.ANALYZE_OURDATA
ADD neg2logpvalusa AS -2 * log(pvalusa) PERSISTED;
GO

ALTER TABLE dbo.ANALYZE_OURDATA
ADD neg2logpvalfra AS -2 * log(pvalfra) PERSISTED;
GO

ALTER TABLE dbo.ANALYZE_OURDATA
ADD neg2logpvalbos AS -2 * log(pvalbos) PERSISTED;
GO

ALTER TABLE dbo.ANALYZE_OURDATA
ADD neg2logpvaluk AS -2 * log(pvaluk) PERSISTED;
GO


---create no key table
CREATE TABLE dbo.[ANALYZE_NOKEY]
([CHROMOSOME] [int] NULL,
      [SNP] [varchar](30) NULL,
      [BP] [int] NULL,
      [PVALUE] [float] NULL,
      [POP] [varchar](50) NULL)
GO


----Previously before haploview method, logs are calculated (no more
necessary)
ALTER TABLE GWA_UK_BOS_HOL
ADD id INT IDENTITY(1,1) PRIMARY KEY;
GO

ALTER TABLE GWA_UK_BOS_HOL
ADD log_holbos_plot AS (-1 * log10(newpvalueholbos)) PERSISTED;
GO

ALTER TABLE GWA_UK_BOS_HOL
ADD log_holbosuk_plot AS (-1 * log10(newpvalueholbosuk)) PERSISTED;
GO

ALTER TABLE GWA_UK_BOS_HOL
ADD log_holuk_plot AS (-1 * log10(newpvalueholuk)) PERSISTED;
GO
```

```
create table GWA_UK_BOS
(chr int,
snp varchar(30),
bp int,
pvalbos float,
pvaluk float,
newpvalueukbos float)
go

BULK INSERT GWA_UK_BOS
FROM '\\vm3\Alsod\DataSets\DataPlots\uk_bos.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE GWA_UK_BOS
ADD id INT IDENTITY(1,1) PRIMARY KEY;
GO

ALTER TABLE GWA_UK_BOS
ADD log_ukbos_plot AS (-1 * log10(newpvalueukbos)) PERSISTED;
GO
```

Appendix 37 – Updating database using T-SQL

```
----TO DELETE
delete from dbo.visitors where
[Page]='http://alsod.iop.kcl.ac.uk/Als/charts/hitcountstatistics.aspx'
go
delete from dbo.visitors where [Country]='UNKNOWN'
go

---YESTERDAY'S DATE
SELECT DATEADD(d,DATEDIFF(d,1,GETDATE()),0) As 'Yesterday'
go

----pages accessed from yesterday
SELECT *
FROM visitors
WHERE DateTime >= DATEADD(d,DATEDIFF(d,1,GETDATE()),0)
AND DateTime < DATEADD(d,DATEDIFF(d,0,GETDATE()),0)
go

---pages by frequency for yesterday
SELECT Country, COUNT(*) AS 'Frequency'
FROM visitors
WHERE [DateTime] BETWEEN DATEADD(d,DATEDIFF(d,1,GETDATE()),0)
AND DATEADD(d,DATEDIFF(d,0,GETDATE()),0)
GROUP BY Country ORDER BY Country

SELECT TOP 1 [BP] FROM [GWA_UK] WHERE ([CHROMOSOME] = 6) ORDER BY [BP]

--- Daily data on visitors
select top (100) [DateTime] as [Date],COUNT(*) AS 'Frequency' from
dbo.visitors group by [DateTime] order by [DateTime]

SELECT TOP (100) LEFT(DateTime, 12) AS Date, COUNT(*) AS 'Frequency'
FROM dbo.visitors
GROUP BY DateTime

SELECT TOP (100) PERCENT LEFT(DateTime, 12) AS Daily, COUNT(*) AS
'Frequency'
FROM dbo.visitors
GROUP BY DateTime
ORDER BY DateTime

select distinct [Date], COUNT(*) AS 'Frequency' from
[Alsod].[KCLAD\spngoka].[daily_visitors_report]
GROUP BY [Date]
ORDER BY [Date]

--Chosen query. works fine but too long.
select distinct [Daily], COUNT(*) AS 'Frequency' from
[Alsod].[KCLAD\spngoka].[daily_visit3]
GROUP BY [Daily]
ORDER BY [Daily]

SELECT DATENAME(mm-yyyy, DateTime) AS Daily FROM dbo.visitors GROUP BY
DATENAME(mm-yyyy, DateTime)

SELECT CONVERT (SMALLDATETIME, CONVERT (VARCHAR, [DateTime],103),103) As
Daily FROM dbo.visitors order by Daily
```

```
SELECT [DateTime](Month,getDate()) + ' ' + [DateTime](Year,getDate()) AS
Daily FROM dbo.visitors GROUP BY Daily

SELECT
    DATENAME(mm, DateTime) AS Month,
    DATENAME(yyyy, DateTime) AS Year,
    COUNT(*) AS Frequency
FROM dbo.visitors AS article
GROUP BY
    DATENAME(mm, DateTime),
    DATENAME(yyyy, DateTime)
ORDER BY Year

--Not quite. unordered!
SELECT
    DATENAME(mm, Daily) AS Month,
    DATENAME(yyyy, Daily) AS Year,
    COUNT(*) AS Frequency
FROM [Alsod].[KCLAD\spngoka].[daily_visit3] AS article
GROUP BY
    DATENAME(mm, Daily),
    DATENAME(yyyy, Daily)

Select Distinct DateName(mm, [DateTime]) as 'Month',
year([DateTime])
as 'Year'from dbo.visitors
Order By DateName(yyyy, [DateTime])ASC



select CONVERT(datetime, [DateTime], 103) as [Date]from dbo.visitors
GROUP BY [DateTime]

SELECT     TOP (100) LEFT(DateTime, 12) AS Date, COUNT(*) AS 'Frequency'
FROM         dbo.visitors
GROUP BY DateTime

SELECT         TOP (100) LEFT([DateTime], 12) AS [Date], COUNT(distinct
[DateTime]) AS 'Frequency', ID
FROM         dbo.visitors
GROUP BY [DateTime], ID
ORDER BY ID


SELECT     TOP (100) Country, COUNT(*) AS 'Frequency'
FROM         dbo.visitors
GROUP BY Country
ORDER BY Country

---Genome table from ucsc webpage
USE [Alsod]
GO

CREATE TABLE [genome](
      [structure_id] [varchar](50) NOT NULL,
      [chrom] [varchar](15) NOT NULL,
      [strand] [varchar](1) NULL,
      [txtStart] [int] NULL,
      [txtEnd] [int] NULL,
      [cdsStart] [int] NULL,
      [cdsEnd] [int] NULL,
      [exonCount] [int] NULL,
```

```
	[exonStarts] [text] NULL,
	[exonEnds] [text] NULL,
	[proteinID] [varchar](50) NULL,
	[alignID] [varchar](50) NULL)
GO

BULK INSERT [genome]
FROM '\\vm3\Alsod\DataSets\genome.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE [genome]
ADD id INT IDENTITY(1,1) PRIMARY KEY;
GO

---Genesequence import
CREATE TABLE [gene_sequence_new](
	[gene_id] [varchar](64) NOT NULL,
	[ncbi_refseq_id] [varchar](50) NULL,
	[strand] [varchar](1) NULL,
	[genomic_size] [int] NULL,
	[exon_count] [int] NULL,
	[coding_exon_count] [int] NULL,
	[structure_id] [varchar](50) NOT NULL,
	[chrom] [int] NOT NULL,
	[rangefrom] [int] NOT NULL,
	[rangeto] [int] NOT NULL,
	[sequence_location_type] [varchar](24) NULL,
	[sequence] [text] NULL)

GO

BULK INSERT [gene_sequence_new]
FROM '\\vm3\Alsod\DataSets\GeneSequence1.txt'
WITH
(
FIELDTERMINATOR = ',',
ROWTERMINATOR = '\n\n'
)
GO

drop table [gene_sequence_new]
go

Exec master.dbo.xp_cmdshell  '\\vm3\Alsod\test.bat',no_output
go

--checking schema
Select * From Information_Schema.Tables
go

--adds user to bulkadmin role
EXEC sp_addsrvrolemember 'KCLAD\spngoka', 'bulkadmin'
go

--Grant permissions
USE Alsod;
GRANT CONTROL ON SCHEMA::dbo TO public
```

```
GO

--To truncate tables
TRUNCATE TABLE dbo.GWA_UK_BOS_FRA_USA
go

--Create table for gene SNPs
CREATE TABLE dbo.[gene_SNP](
      [gene_id] [varchar](64) PRIMARY KEY,
      [snp] [varchar](50) NOT NULL,
      [basepair] [int] NOT NULL,
      [pvalue] [float] NULL,
      [pubmed_id] [varchar](20) NULL)

GO

ALTER TABLE dbo.[gene_SNP]
ADD [term] [varchar](10) NULL;
GO
ALTER TABLE dbo.[gene_SNP]
ADD S_No int IDENTITY(1,1);
GO
ALTER TABLE dbo.[gene_SNP]
ADD [First_Author] [varchar](50) NULL;
GO
ALTER TABLE dbo.[gene_SNP]
ADD [Year] [varchar](50) NULL;
GO

---An updated gene SNPlist
CREATE TABLE dbo.[gene_SNP_updated](
      [gene_id] [varchar](64) NULL,
      [snp] [varchar](50) NULL,
      [basepair] [int] NULL,
      [pvalue] [float] NULL,
      [pubmed_id] [varchar](20) NULL,
      [paperlink] [varchar](1024)NULL,
      [term] [varchar](10) NULL,
      [First_Author] [varchar](50) NULL,
      [Year] [varchar](50) NULL,
      [link] [varchar](1024) NULL,
      [Amino_acid] [varchar](50) NULL)
GO

BULK INSERT dbo.[gene_SNP_updated]
FROM '\\vm3\Alsod\SNP\updated_snps.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.[gene_SNP_updated]
ADD S_No int IDENTITY(1,1);
GO



---Create table for gene SNPs from ALSGene site

CREATE TABLE dbo.[gene_SNP_ALSGene](
```

```
	[chromosome_name] [varchar](4) NOT NULL,
	[gene_id] [varchar](64) NOT NULL,
	[ncbi_locuslink_id] int NULL,
	[gene_name] [varchar](1024) NULL,
	[First_Author] [varchar](50) NULL,
	[Year] [varchar](50) NULL,
	[pubmed_id] [varchar](20) NULL,
	[snp1] [varchar](50) NULL,
	[snp2] [varchar](50) NULL,
	[mutation_mnemonic] [varchar](1024) NULL,
	[codon] int NULL,
	[paperlink] [varchar](1024)NULL,
	[basepair] [int] NULL,
	[term] [varchar](10) NULL,
	[pvalue] float NULL)
GO


BULK INSERT dbo.[gene_SNP_ALSGene]
FROM '\\vm3\Alsod\SNP\split_ALSGenesnps.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO


UPDATE dbo.[gene] SET [ALSgene_id] = '106' where [keywords] ='ALS2'
GO


---Extras after bulk insert
ALTER TABLE dbo.[gene_SNP_ALSGene]
ADD [S_No] [int] IDENTITY(1,1);
GO

UPDATE dbo.[gene_SNP_ALSGene] SET [term] = 'GRCh37'
GO

UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association of
apolipoprotein E epsilon 4 allele with bulbar-onset motor neuron disease'
WHERE [First_Author] = 'Al-Chalabi'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Lack of association of
apolipoprotein E epsilon 4 allele with bulbar-onset motor neuron disease'
WHERE [First_Author] = 'Bachus'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'The P413L chromogranin B
variation in French patients with sporadic amyotrophic lateral sclerosis'
WHERE [First_Author] = 'Blasco' AND [Year] = '2010'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association between
divalent metal transport 1 encoding gene (SLC11A2) and disease duration in
amyotrophic lateral sclerosis' WHERE [First_Author] = 'Blasco' AND [Year] =
'2011'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Polymorphisms in the GluR2
gene are not associated with amyotrophic lateral sclerosis' WHERE
[First_Author] = 'Bogaert'
GO
```

```
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Screening of the regulatory and coding regions of vascular endothelial growth factor in amyotrophic lateral sclerosis' WHERE [First_Author] = 'Brockington' AND [Year] = '2005'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Screening of the transcriptional regulatory regions of vascular endothelial growth factor receptor 2 (VEGFR2) in amyotrophic lateral sclerosis' WHERE [First_Author] = 'Brockington' AND [Year] = '2007'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Increased incidence of FMO1 gene single nucleotide polymorphisms in sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Cereda' AND [Year] = '2006'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'TNF and sTNFR1/2 plasma levels in ALS patients' WHERE [First_Author] = 'Cereda' AND [Year] = '2008'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Lack of association of VEGF promoter polymorphisms with sporadic ALS' WHERE [First_Author] = 'Chen'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'A two-stage genome-wide association study of sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Chio'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Sporadic ALS is not associated with VAPB gene mutations in Southern Italy' WHERE [First_Author] = 'Conforti' AND [Year] = '2006'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'A novel Angiogenin gene mutation in a sporadic patient with amyotrophic lateral sclerosis from southern Italy' WHERE [First_Author] = 'Conforti' AND [Year] = '2008'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association of the hOGG1 Ser326Cys polymorphism with sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Coppede' AND [Year] = '2007'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Lack of association between the APEX1 Asp148Glu polymorphism and sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Coppede' AND [Year] = '2010'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association study between XRCC1 gene polymorphisms and sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Coppede' AND [Year] = '2010'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Variations in the coding and regulatory sequences of the angiogenin (ANG) gene are not associated to ALS (amyotrophic lateral sclerosis) in the Italian population' WHERE [First_Author] = 'Corrado' AND [Year] = '2007'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'VPS54 genetic analysis in ALS Italian cohort' WHERE [First_Author] = 'Corrado' AND [Year] = '2010'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Paraoxonase promoter and intronic variants modify risk of sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Cronin' AND [Year] = '2007'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'A genome-wide association study of sporadic ALS in a homogenous Irish population' WHERE [First_Author] = 'Cronin' AND [Year] = '2008'
GO
```

```
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Analysis of the UNC13A gene as a risk factor for sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Daoud'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Absence of angiogenic genes modification in Italian ALS patients' WHERE [First_Author] = 'Del Bo' AND [Year] = '2008'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'No major progranulin genetic variability contribution to disease etiopathogenesis in an ALS Italian cohort' WHERE [First_Author] = 'Del Bo' AND [Year] = '2009'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association of APOE epsilon4 allele with survival in amyotrophic lateral sclerosis' WHERE [First_Author] = 'Drory'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Whole-genome analysis of sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Dunckley'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS' WHERE [First_Author] = 'Elden'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'No association between DNA repair gene XRCC1 and amyotrophic lateral sclerosis' WHERE [First_Author] = 'Fang'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Possible gender-dependent association of vascular endothelial growth factor (VEGF) gene and ALS' WHERE [First_Author] = 'Fernandez-Santiago'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Apolipoprotein E and multiple sclerosis: a biochemical and genetic investigation' WHERE [First_Author] = 'Gaillard'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Androgen receptor gene polymorphisms in amyotrophic lateral sclerosis' WHERE [First_Author] = 'Garofalo'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Identification of new ANG gene mutations in a large cohort of Italian patients with amyotrophic lateral sclerosis' WHERE [First_Author] = 'Gellera'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Is erythropoietin gene a modifier factor in amyotrophic lateral sclerosis?' WHERE [First_Author] = 'Ghezzi'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Lack of association between VEGF gene polymorphisms and plasma VEGF levels and sporadic AL' WHERE [First_Author] = 'Golenia'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association of the H63D polymorphism in the hemochromatosis gene with sporadic ALS' WHERE [First_Author] = 'Goodall'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'A novel candidate region for ALS on chromosome 14q11.2' WHERE [First_Author] = 'Greenway' AND [Year] = '2004'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis' WHERE [First_Author] = 'Greenway' AND [Year] = '2006'
```

```
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'TDP-43 is not a common cause of sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Guerreiro'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Molecular genetic analysis of the APEX nuclease gene in amyotrophic lateral sclerosis' WHERE [First_Author] = 'Hayward'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'H63D polymorphism in the hemochromatosis gene is associated with sporadic amyotrophic lateral sclerosis in China' WHERE [First_Author] = 'He'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Replication analysis of SNPs on 9p21.2 and 19p13.3 with amyotrophic lateral sclerosis in East Asians' WHERE [First_Author] = 'Iida'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Amyotrophic lateral sclerosis, lead, and genetic susceptibility: polymorphisms in the delta-aminolevulinic acid dehydratase and vitamin D receptor genes' WHERE [First_Author] = 'Kamel'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Large-scale pathways-based association study in amyotrophic lateral sclerosis' WHERE [First_Author] = 'Kasperaviciute'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'The thermolabile variant of 5,10-methylenetetrahydrofolate reductase is a possible risk factor for amyotrophic lateral sclerosis' WHERE [First_Author] = 'Kuehnlein'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'APOE: a potential marker of disease progression in ALS' WHERE [First_Author] = 'Laaksovirta'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motoneurons against ischemic death' WHERE [First_Author] = 'Lacomblez'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'A common haplotype within the PON1 promoter region is associated with sporadic ALS' WHERE [First_Author] = 'Lambrechts'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Landers' AND [Year] = '2008'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Ataxin-2 intermediate-length polyglutamine expansions in European ALS patients' WHERE [First_Author] = 'Landers' AND [Year] = '2009'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Prion protein codon 129 genotype prevalence is altered in primary progressive aphasia' WHERE [First_Author] = 'Lee'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Evaluation of the Golgi trafficking protein VPS54 (wobbler) as a candidate for ALS' WHERE [First_Author] = 'Li'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Thioredoxin reductase 1 haplotypes modify familial amyotrophic lateral sclerosis onset' WHERE [First_Author] = 'Meisler'
GO
```

```
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'A gene-environment study
of the paraoxonase 1 gene and pesticides in amyotrophic lateral sclerosis'
WHERE [First_Author] = 'Mitchell'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Apolipoprotein E
genotyping in sporadic amyotrophic lateral sclerosis: evidence for a major
influence on the clinical presentation and prognosis' WHERE [First_Author]
= 'Morahan'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Apolipoprotein E
genotyping in sporadic amyotrophic lateral sclerosis: evidence for a major
influence on the clinical presentation and prognosis' WHERE [First_Author]
= 'Moulard'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Apolipoprotein E epsilon 4
allele is not associated with earlier age at onset in amyotrophic lateral
sclerosis' WHERE [First_Author] = 'Mui'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Motoneuron-specific NR3B
gene: no association with ALS and evidence for a common null allele' WHERE
[First_Author] = 'Niemann'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Genotyping of presenilin-1
polymorphism in amyotrophic lateral sclerosis' WHERE [First_Author] =
'Panas'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Spinobulbar muscular
atrophy can mimic ALS: the importance of genetic testing in male patients
with atypical ALS' WHERE [First_Author] = 'Parboosingh'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Mutations of the ANG gene
in French patients with sporadic amyotrophic lateral sclerosis' WHERE
[First_Author] = 'Paubel'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'New application of
intelligent agents in sporadic amyotrophic lateral sclerosis identifies
unexpected specific genetic background' WHERE [First_Author] = 'Penco'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'HFE H63D polymorphism is
increased in patients with amyotrophic lateral sclerosis of Italian origin'
WHERE [First_Author] = 'Restagno'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Lack of association of PON
polymorphisms with sporadic ALS in an Italian population' WHERE
[First_Author] = 'Ricci'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Paraoxonase cluster
polymorphisms are associated with sporadic ALS' WHERE [First_Author] =
'Saeed'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'A polymorphism in the
poliovirus receptor gene differs in motor neuron disease' WHERE
[First_Author] = 'Saunderson'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association of ALS with
head injury, cigarette smoking and APOE genotypes' WHERE [First_Author] =
'Schmidt'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'No association of DYNC1H1
with sporadic ALS in a case-control study of a northern European derived
population: a tagging SNP approach' WHERE [First_Author] = 'Shah'
GO
```

```
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study' WHERE [First_Author] = 'Shatunov'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Lack of association between apolipoprotein E genotype and sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Siddique'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Analysis of heavy neurofilament subunit gene polymorphism in Russian patients with sporadic motor neuron disease (MND)' WHERE [First_Author] = 'Skvortsova'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Paraoxonase gene polymorphisms and sporadic ALS' WHERE [First_Author] = 'Slowik'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'The association between H63D mutations in HFE and amyotrophic lateral sclerosis in a Dutch population' WHERE [First_Author] = 'Sutedja'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Development of a high-throughput microarray-based resequencing system for neurological disorders and its application to molecular genetics of amyotrophic lateral sclerosis' WHERE [First_Author] = 'Takahashi'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'VEGF promoter haplotype and amyotrophic lateral sclerosis (ALS)' WHERE [First_Author] = 'Terry'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Distribution of MnSOD polymorphisms in sporadic ALS patients' WHERE [First_Author] = 'Tomblyn'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Screening of AP endonuclease as a candidate gene for amyotrophic lateral sclerosis (ALS)' WHERE [First_Author] = 'Tomkins' AND [Year] = '2000'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Mutation screening of manganese superoxide dismutase in amyotrophic lateral sclerosis' WHERE [First_Author] = 'Tomkins' AND [Year] = '2001'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association of paraoxonase gene cluster polymorphisms with ALS in France, Quebec, and Sweden' WHERE [First_Author] = 'Valdmanis'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study' WHERE [First_Author] = 'van Es' AND [Year] = '2007'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis' WHERE [First_Author] = 'van Es' AND [Year] = '2008'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'van Es' AND [Year] = '2009'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Manganese-containing superoxide dismutase signal sequence polymorphism associated with sporadic motor neuron disease' WHERE [First_Author] = 'Van Lendeghem'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Lack of association between VEGF polymorphisms and ALS in a Dutch population' WHERE [First_Author] = 'Van Vught' AND [Year] = '2005'
```

```
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'P413L CHGB is not associated with ALS susceptibility or age at onset in a Dutch population' WHERE [First_Author] = 'van Vught' AND [Year] = '2010'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Risk genotypes at TMEM106B are associated with cognitive impairment in amyotrophic lateral sclerosis' WHERE [First_Author] = 'Vass'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Increased incidence of the Hfe mutation in amyotrophic lateral sclerosis and related cellular consequences' WHERE [First_Author] = 'Wang'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Genetic analysis of the cystatin C gene in familial and sporadic ALS patients' WHERE [First_Author] = 'Watanabe'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Paraoxonase 1 (PON1) organophosphate hydrolysis is not reduced in ALS' WHERE [First_Author] = 'Wills'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Quantification of cystatin C in cerebrospinal fluid from various neurological disorders and correlation with G73A polymorphism in CST3' WHERE [First_Author] = 'Yamamoto-Watanabe'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'HFE mutations are not strongly associated with sporadic ALS' WHERE [First_Author] = 'Yen'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'The C(-1562)T polymorphism of the MMP-9 gene and the risk of sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Zawislak' AND [Year] = '2009'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'The -A162G polymorphism of the PON1 gene and the risk of sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Zawislak' AND [Year] = '2010'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'Association of APOE with age at onset of sporadic amyotrophic lateral sclerosis' WHERE [First_Author] = 'Zetterberg'
GO
UPDATE dbo.[gene_SNP_ALSGene] SET [paperlink] = 'VEGF C2578A polymorphism does not contribute to amyotrophic lateral sclerosis susceptibility in sporadic Chinese patients' WHERE [First_Author] = 'Zhang'
GO

---Create table for gene SNPs from lifeseq site

CREATE TABLE dbo.[gene_SNP_lifeseq](
      [gene_id] [varchar](500) NOT NULL,
      [DNA_change_original_from] [varchar](200) NULL,
      [DNA_change_original_to] [varchar](200) NULL,
      [chromosome_name] [varchar](4) NULL,
      [position] [varchar](100) NULL,
      [basepair] [int] NULL,
      [exon] [varchar](100) NULL,
      [snp2] [varchar](500) NULL,
      [mutation_mnemonic] [varchar](1024) NULL,
      [codon] int NULL,
      [mRNA_accession_no] [varchar](200) NULL,
      [amino_acid_accession_no] [varchar](200) NULL,
      [amino_acid_change_as_seen] [varchar](1024) NULL,
```

```
	[zygosity] [varchar](200) NULL,
	[ethnicity] [varchar](500) NULL,
	[country] [varchar](500) NULL,
	[families] [varchar](100) NULL,
	[number_patients_with_mutations] [varchar](100) NULL,
	[rate_patients_with_mutation] [varchar](100) NULL,
	[rate_patients] [varchar](100) NULL,
	[rate_patients_mutation_found] int NULL,
	[rate_patients_mutation_not_found] int NULL,
	[rate_controls] [varchar](100) NULL,
	[rate_controls_mutation_found] int NULL,
	[rate_controls_mutation_not found] int NULL,
	[odd_ratio] [varchar](100) NULL,
	[CI(95%)] [varchar](200) NULL,
	[Chi-Square_p_value] [varchar](100) NULL,
	[clinical_characteristics] [text] NULL,
	[family_history] [varchar](200) NULL,
	[gender] [varchar](100) NULL,
	[age_onset] [varchar](100) NULL,
	[duration_in_text] [varchar](100) NULL,
	[duration_months] [varchar](100) NULL,
	[UMN] [varchar](100) NULL,
	[LMN] [varchar](100) NULL,
	[Site_onset] [varchar](100) NULL,
	[years_until_initiation_respirator] [varchar](100) NULL,
	[pubmed_id] [varchar](20) NULL,
	[paper_title] [varchar](1024) NULL,
	[First_Author] [varchar](50) NULL,
	[Year] [varchar](50) NULL,
	[comment] [text] NULL,
	[journal] [varchar](1024) NULL)

GO

BULK INSERT dbo.[gene_SNP_lifeseq]
FROM '\\vm3\Alsod\SNP\split_lifeseq.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO



---A catalogue of GWAS
create table dbo.GWA_CATALOGUE(
Added_date varchar(250) NULL,
Pubmed_id int NULL,
First_author varchar(250),
Submission_date varchar(250) NULL,
Journal varchar(1024) NULL,
Link varchar(1024) NULL,
Study varchar(500) NULL,
Disease_trait varchar(1024) NULL,
Initial_sample_size varchar(1024) NULL,
Replication_sample_size varchar(1024) NULL,
Region varchar(500) NULL,
Reported_genes varchar(250) NULL,
Strongest_snps_risk_allele varchar(250) NULL,
Snps varchar(250) NULL,
Risk_allele_frequency varchar(250) NULL,
```

```
P_value varchar(250) NULL,
P_value_text varchar(250) NULL,
Odd_ratio_beta varchar(200) NULL,
CI_95_percent_text varchar(200) NULL,
Platform_snp_passing_QC varchar(500) NULL,
CNV varchar(100) NULL)
go

--originally from http://www.genome.gov/admin/gwascatalog.txt
BULK INSERT dbo.GWA_CATALOGUE
FROM '\\vm3\Alsod\SNP\gwascatalog.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.GWA_CATALOGUE
ADD S_No int IDENTITY(1,1) PRIMARY KEY;
GO

--- To search for rows with Amyotrophic Lateral Sclerosis
SELECT      Added_date, Pubmed_id, First_author, Submission_date, Journal,
Link, Study, Disease_trait, Initial_sample_size, Replication_sample_size,
Region,
                    Reported_genes,   Strongest_snps_risk_allele,   Snps,
Risk_allele_frequency,      P_value,      P_value_text,      Odd_ratio_beta,
CI_95_percent_text,
                    Platform_snp_passing_QC, CNV, S_No
FROM         dbo.GWA_CATALOGUE
WHERE     (Disease_trait = 'Amyotrophic Lateral Sclerosis')

----Update date colmn with trimmed characters
UPDATE  dbo.GWA_CATALOGUE  SET  Submission_date  =  RIGHT  (Submission_date,
LEN(Submission_date)-6)
WHERE  (Disease_trait = 'Amyotrophic Lateral Sclerosis')
GO

UPDATE  dbo.GWA_CATALOGUE  SET  Strongest_snps_risk_allele  =  LEFT(REVERSE
(Strongest_snps_risk_allele),1)
WHERE      (Disease_trait   =   'Amyotrophic   Lateral   Sclerosis')   and
(Strongest_snps_risk_allele <>'NR')
GO

----Getting count of snp in gwas als
select count(*) AS 'Total' FROM  dbo.GWA_CATALOGUE WHERE (Disease_trait =
'Amyotrophic Lateral Sclerosis')

---Remove characters from column
select RIGHT (Strongest_snps_risk_allele, LEN(Strongest_snps_risk_allele)-
1)    as    Strongest_snps_risk_allele    from    dbo.GWA_CATALOGUE    WHERE
(Disease_trait      =      'Amyotrophic      Lateral      Sclerosis')      and
(Strongest_snps_risk_allele <>'NR')

select        LEFT(REVERSE        (Strongest_snps_risk_allele),1)        as
Strongest_snps_risk_allele from dbo.GWA_CATALOGUE WHERE  (Disease_trait =
'Amyotrophic Lateral Sclerosis') and (Strongest_snps_risk_allele <>'NR')

----Alter gene table
ALTER TABLE dbo.gene
ADD [snp] [varchar](50) NULL;
```

```
GO
ALTER TABLE dbo.gene
ADD [basepair] [int] NULL;
GO
ALTER TABLE dbo.gene
ADD [pvalue] [float] NULL;
GO
ALTER TABLE dbo.gene
ADD [snp_pubmed_id] [varchar](20) NULL;
GO
ALTER TABLE dbo.gene
ADD [ALSgene_id] [varchar](10) NULL;
GO

ALTER TABLE dbo.gene
ADD [snp_paperlink] [varchar](1024) NULL;
GO


ALTER TABLE dbo.gene
ADD [lifesciencedb] [varchar](10) NULL;
GO

UPDATE dbo.gene SET keywords = 'NEFH'
WHERE gene_id = 'NEFH'
GO

UPDATE dbo.patient_record SET country_iso_code = 'BR'
WHERE patient_id = '2215'
GO

UPDATE dbo.gene SET graph = '1'
GO

SELECT      Pubmed_id,      First_author,      Submission_date,      Journal,
Initial_sample_size,   Replication_sample_size,   Region,   Reported_genes,
Strongest_snps_risk_allele,     Snps,     Risk_allele_frequency,     P_value,
P_value_text,  Odd_ratio_beta,  CI_95_percent_text,  Platform_snp_passing_QC,
CNV,   S_No,   Added_date   FROM   dbo.GWA_CATALOGUE   WHERE   (Disease_trait   =
'Amyotrophic Lateral Sclerosis')
GO

UPDATE dbo.gene SET dbSNP = 'rs74315431'
WHERE gene_id = 'VAPB'
GO

UPDATE dbo.gene SET pubmed_id1 = '15060112'
WHERE gene_id = 'VAPB'
GO

---Get ALSGene id for link to alsgene.org
UPDATE dbo.gene SET ALSgene_id = '67' WHERE gene_id = 'ALAD'
GO
UPDATE dbo.gene SET ALSgene_id = '1' WHERE gene_id = 'ALS2'
GO
UPDATE dbo.gene SET ALSgene_id = '7' WHERE gene_id = 'ANG'
GO
UPDATE dbo.gene SET ALSgene_id = '2' WHERE gene_id = 'APEX1'
GO
UPDATE dbo.gene SET ALSgene_id = '25' WHERE gene_id = 'APOE'
GO
UPDATE dbo.gene SET ALSgene_id = '82' WHERE gene_id = 'AR'
```

```
GO
UPDATE dbo.gene SET ALSgene_id = '53' WHERE gene_id = 'ATXN2'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'B4GALT6'
GO
UPDATE dbo.gene SET ALSgene_id = '16' WHERE gene_id = 'C9orf72'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'CCS'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'CHMP2B'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'CNTF'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'CNTN4'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'CRYM'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'CSNK1G3'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'CYP2D6'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'DAO'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'DCTN1'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'DISC1'
GO
UPDATE dbo.gene SET ALSgene_id = '12' WHERE gene_id = 'DPP6'
GO
UPDATE dbo.gene SET ALSgene_id = '88' WHERE gene_id = 'DYNC1H1'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'EFEMP1'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'ELP3'
GO
UPDATE dbo.gene SET ALSgene_id = '10' WHERE gene_id = 'FGGY'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'FIG4'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'FUS'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'GARS'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'HEXA'
GO
UPDATE dbo.gene SET ALSgene_id = '8' WHERE gene_id = 'HFE'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'IFNK'
GO
UPDATE dbo.gene SET ALSgene_id = '11' WHERE gene_id = 'ITPR2'
GO
UPDATE dbo.gene SET ALSgene_id = '14' WHERE gene_id = 'KIFAP3'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'LIF'
GO
UPDATE dbo.gene SET ALSgene_id = '37' WHERE gene_id = 'LIPC'
GO
UPDATE dbo.gene SET ALSgene_id = '1' WHERE gene_id = 'LOX'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'LUM'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'MAOB'
```

```
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'MAPT'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'MOBKL2B'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'MT-ND2'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'NAIP'
GO
UPDATE dbo.gene SET ALSgene_id = '43' WHERE gene_id = 'NEFH'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'NT5C1A'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'OPTN'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'PGRN'
GO
UPDATE dbo.gene SET ALSgene_id = '3' WHERE gene_id = 'PON1'
GO
UPDATE dbo.gene SET ALSgene_id = '4' WHERE gene_id = 'PON2'
GO
UPDATE dbo.gene SET ALSgene_id = '5' WHERE gene_id = 'PON3'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'PRPH'
GO
UPDATE dbo.gene SET ALSgene_id = '73' WHERE gene_id = 'PSEN1'
GO
UPDATE dbo.gene SET ALSgene_id = '72' WHERE gene_id = 'PVR'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'RBMS1'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SCN7A'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SELL'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SEMA6A'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SETX'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SLC1A2'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SLC39A11'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SMN1'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SMN2'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SNCG'
GO
UPDATE dbo.gene SET ALSgene_id = '1' WHERE gene_id = 'SOD1'
GO
UPDATE dbo.gene SET ALSgene_id = '50' WHERE gene_id = 'SOD2'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SPAST'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SPG7'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'SUSD1'
GO
UPDATE dbo.gene SET ALSgene_id = '77' WHERE gene_id = 'TARDBP(TDP43)'
GO
UPDATE dbo.gene SET ALSgene_id = '52' WHERE gene_id = 'UNC13A'
```

```
GO
UPDATE dbo.gene SET ALSgene_id = '69' WHERE gene_id = 'VAPB'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'VCP'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'VDR'
GO
UPDATE dbo.gene SET ALSgene_id = '9' WHERE gene_id = 'VEGFA'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'ZFP64'
GO
UPDATE dbo.gene SET ALSgene_id = '0' WHERE gene_id = 'ZNF746'
GO

----Create table for gene SNPs
CREATE TABLE dbo.[mutation_replication](
      [gene] [varchar](30) NOT NULL,
      [mutation] [varchar](100) NULL,
      [codon] [varchar](10) NULL,
      [families] [varchar](10) NULL,
      [generations] [varchar](10) NULL,
      [affected_patients] [varchar](10) NULL,
      [sporadic] [varchar](10) NULL,
      [familial] [varchar](10) NULL,
      [family_history] [varchar](50) NULL,
      [cases] int NULL,
      [controls] int NULL,
      [score] float NULL,
      [ethnicity] [varchar](500) NULL,
      [country1] [varchar](500) NULL,
      [country2] [varchar](500) NULL,
      [country3] [varchar](500) NULL,
      [country4] [varchar](500) NULL,
      [country5] [varchar](500) NULL,
      [snp] [varchar](500) NULL,
      [pathogenic] [varchar](10) NULL,
      [author] [varchar](50) NULL,
      [year] [varchar](10) NULL,
      [pubmed_id] int NULL,
      [google] int NULL,
      [pubmed] int NULL,
      [journal] int NULL,
      [controversies] [varchar](500) NULL,
      [comment] [varchar](500)NULL)
GO

BULK INSERT dbo.[mutation_replication]
FROM '\\vm3\Alsod\SNP\mutation_replication.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

ALTER TABLE dbo.[mutation_replication]
ADD S_No int IDENTITY(1,1);
GO

ALTER TABLE dbo.[mutation_replication]
ADD [title] [varchar](1024) NULL;
GO
```

```
ALTER TABLE dbo.[mutation_replication]
ADD [link] [varchar](1024) NULL;
GO

ALTER TABLE dbo.[mutation_replication]
ADD [country6] [varchar](500) NULL;
GO

ALTER TABLE dbo.[mutation_replication]
ADD [country7] [varchar](500) NULL;
GO

ALTER TABLE dbo.[mutation_replication]
ADD [country8] [varchar](500) NULL;
GO

ALTER TABLE dbo.[mutation_replication]
ADD [country9] [varchar](500) NULL;
GO

ALTER TABLE dbo.[mutation_replication]
ADD [country10] [varchar](500) NULL;
GO

ALTER TABLE dbo.gene
ADD [inheritance_pattern] [varchar](1024) NULL;
GO

UPDATE dbo.gene SET inheritance_pattern = 'autosomal dominant' WHERE
gene_id = 'ANG'
GO
UPDATE dbo.gene SET inheritance_pattern = 'dominant and recessive' WHERE
gene_id = 'SOD1'
GO
UPDATE dbo.gene SET inheritance_pattern = 'autosomal recessive' WHERE
gene_id = 'ALS2'
GO
UPDATE dbo.gene SET inheritance_pattern = 'unknown' WHERE gene_id = 'NEFH'
GO
UPDATE dbo.gene SET inheritance_pattern = 'classic adult-onset ' WHERE
gene_id = 'DAO'
GO

--Hereditary motor syndromes IDs
ALTER TABLE dbo.[gene]
ADD [syndrome_id] [varchar](50) NULL;
GO
UPDATE dbo.gene SET syndrome_id = 'so' WHERE gene_id = 'SOD1'
GO
UPDATE dbo.gene SET syndrome_id = 'als4' WHERE gene_id = 'SETX'
GO
UPDATE dbo.gene SET syndrome_id = 'als16q' WHERE gene_id = 'FUS'
GO
UPDATE dbo.gene SET syndrome_id = 'als20q' WHERE gene_id = 'VAPB'
GO
UPDATE dbo.gene SET syndrome_id = 'alsang' WHERE gene_id = 'ANG'
GO
UPDATE dbo.gene SET syndrome_id = 'tdp43als' WHERE gene_id =
'TARDBP(TDP43)'
GO
UPDATE dbo.gene SET syndrome_id = 'alsfig4f' WHERE gene_id = 'FIG4'
```

```
GO
UPDATE dbo.gene SET syndrome_id = 'alsoptn' WHERE gene_id = 'OPTN'
GO
UPDATE dbo.gene SET syndrome_id = 'sca2' WHERE gene_id = 'ATXN2'
GO
UPDATE dbo.gene SET syndrome_id = 'vcpals' WHERE gene_id = 'VCP'
GO
UPDATE dbo.gene SET syndrome_id = 'daoals' WHERE gene_id = 'DAO'
GO
UPDATE dbo.gene SET syndrome_id = 'alsftd2' WHERE gene_id = 'C9orf72'
GO
UPDATE dbo.gene SET syndrome_id = 'vocaldyn' WHERE gene_id = 'DCTN1'
GO
UPDATE dbo.gene SET syndrome_id = 'cals' WHERE gene_id = 'ALS2'
GO
UPDATE dbo.gene SET syndrome_id = 'als5' WHERE gene_id = 'SPAST'
GO
UPDATE dbo.gene SET syndrome_id = 'alssigmar1' WHERE gene_id = 'SIGMAR1'
GO
UPDATE dbo.gene SET syndrome_id = 'alsx' WHERE gene_id = 'UBQLN2'
GO
UPDATE dbo.gene SET syndrome_id = 'alsftd3' WHERE gene_id = 'CHMP2B'
GO
UPDATE dbo.gene SET syndrome_id = 'nf' WHERE gene_id = 'NEFH'
GO
UPDATE dbo.gene SET syndrome_id = 'peripherinals' WHERE gene_id = 'PRPH'
GO
UPDATE dbo.gene SET syndrome_id = 'taf15als' WHERE gene_id = 'TAF15'
GO
UPDATE dbo.gene SET syndrome_id = 'nil' WHERE syndrome_id = NULL
GO

--Causative flag
ALTER TABLE dbo.[gene]
ADD [causative_id] [varchar](50) NULL;
GO
UPDATE dbo.gene SET causative_id = 'ALS 1' WHERE gene_id = 'SOD1'
GO
UPDATE dbo.gene SET causative_id = 'ALS 4' WHERE gene_id = 'SETX'
GO
UPDATE dbo.gene SET causative_id = 'ALS 6' WHERE gene_id = 'FUS'
GO
UPDATE dbo.gene SET causative_id = 'ALS 8' WHERE gene_id = 'VAPB'
GO
UPDATE dbo.gene SET causative_id = 'ALS 9' WHERE gene_id = 'ANG'
GO
UPDATE dbo.gene SET causative_id = 'ALS 10' WHERE gene_id = 'TARDBP(TDP43)'
GO
UPDATE dbo.gene SET causative_id = 'ALS 11' WHERE gene_id = 'FIG4'
GO
UPDATE dbo.gene SET causative_id = 'ALS 12' WHERE gene_id = 'OPTN'
GO
UPDATE dbo.gene SET causative_id = 'ALS 13' WHERE gene_id = 'ATXN2'
GO
UPDATE dbo.gene SET causative_id = 'ALS 14' WHERE gene_id = 'VCP'
GO
UPDATE dbo.gene SET causative_id = 'ALS' WHERE gene_id = 'DAO'
GO
UPDATE dbo.gene SET causative_id = 'ALS-FTD 2' WHERE gene_id = 'C9orf72'
GO
UPDATE dbo.gene SET causative_id = 'ALS' WHERE gene_id = 'DCTN1'
GO
```

```
UPDATE dbo.gene SET causative_id = 'ALS 2' WHERE gene_id = 'ALS2'
GO
UPDATE dbo.gene SET causative_id = 'ALS 5' WHERE gene_id = 'SPAST'
GO
UPDATE dbo.gene SET causative_id = 'ALS 16' WHERE gene_id = 'SIGMAR1'
GO
UPDATE dbo.gene SET causative_id = 'ALS 15' WHERE gene_id = 'UBQLN2'
GO
UPDATE dbo.gene SET causative_id = 'ALS-FTD 3' WHERE gene_id = 'CHMP2B'
GO
UPDATE dbo.gene SET causative_id = 'ALS' WHERE gene_id = 'NEFH'
GO
UPDATE dbo.gene SET causative_id = 'ALS' WHERE gene_id = 'PRPH'
GO
UPDATE dbo.gene SET causative_id = 'ALS' WHERE gene_id = 'TAF15'
GO


---Create frequency table
CREATE TABLE dbo.[gene_frequency](
      [s_no] [int] IDENTITY(1,1),
      [gene] [varchar](30)NULL,
      [country_iso_code] [varchar](2) NULL,
      [replication_id] [int] NULL,
      [Total_cases] [int] NULL,
      [Total_affected] [int] NULL,
      [FALS_number] [int] NULL,
      [FALS_percentage] [float] NULL,
      [SALS_number] [int] NULL,
      [SALS_percentage] [float] NULL,
      [phenotype] [varchar] (150) NULL,
      [comment] [varchar] (1500))
GO

SELECT DISTINCT [S_No], [gene],[author]+' ('+ [year]+')' +' - pubmedID:'+
[pubmed_id] AS [details] FROM [mutation_replication]
GO

ALTER TABLE dbo.patient_record
ADD [phenotype] [varchar](500) NULL;
GO

UPDATE dbo.patient_record SET phenotype = 'ALS'
GO

ALTER TABLE dbo.gene
ADD [FALS_only] [varchar](1) NULL;
GO

UPDATE dbo.gene SET FALS_only = 'Y' WHERE gene_effect = 'FALS genes' OR
gene_effect= 'FALS genes found in SALS'
GO
UPDATE dbo.gene SET FALS_only = 'N' WHERE gene_effect <> 'FALS genes'
GO

SELECT [keywords] FROM dbo.gene
go

---REQUEST BY KARAN LUND (CHRIS SHAW'S LAB 11/04/2013)
SELECT        TOP (100) PERCENT p.sex, c.country_name, p.ethnic_origin,
p.age_of_onset, p.site_of_onset, g.gene_id AS 'Gene', p.alive, g.screened,
```

```
                        g.family_history,     g.gene_id,     m.mutation_type,
m.mutation_name          AS          'Mutation',          m.mutation_mnemonic,
dbo.institution.institution_name,
                        dbo.institution.department,
dbo.institution.contact_first_name,      dbo.institution.contact_last_name,
p.family_id AS 'Author_Year'
FROM          dbo.mutation AS m INNER JOIN
                        dbo.patient_genetic_record  AS  g  ON  m.mutation_id  =
g.mutation_id INNER JOIN
                        dbo.patient_record    AS    p    ON    g.patient_id    =
p.patient_id INNER JOIN
                        dbo.Country    AS    c    ON    p.country_iso_code    =
c.country_iso_code INNER JOIN
                        dbo.institution      ON      m.institution_code      =
dbo.institution.institution_code
WHERE      (m.mutation_id NOT LIKE '%none%') AND (p.age_of_onset <> ' ') AND
(g.gene_id = 'SOD1')
ORDER BY 'Gene', m.codon_id


----updated mutation table created and populated with data 06/05/2013
CREATE TABLE [dbo].[mutation_new](
      [mutation_id] [int] NOT NULL,
      [mutation_alias] [varchar](1024) NULL,
      [mutation_name] [varchar](1024) NOT NULL,
      [mutation_mnemonic] [varchar](1024) NULL,
      [gene_id] [varchar](64) NULL,
      [institution_code] [varchar](36) NULL,
      [mutation_type] [varchar](12) NULL,
      [seq_location_id] [real] NULL,
      [position] [int] NULL,
      [sequence_position_relative] [int] NULL,
      [codon] [int] NULL,
      [codon_id] [int] NULL,
      [sequence_original] [varchar](1024) NULL,
      [sequence_mutated] [varchar](1024) NULL,
      [sequence_location_type] [varchar](50) NULL,
      [sequence_location_number] [int] NULL,
      [aa_original] [varchar](6) NULL,
      [aa_mutated] [varchar](6) NULL,
      [restriction_site] [varchar](7) NULL,
      [enzyme] [varchar](20) NULL,
      [mutation_approved_status] [varchar](128) NULL,
      [mutation_approved_date] [datetime] NULL,
      [mutation_approved_by] [varchar](128) NULL,
      [mutation_submitted_date] [datetime] NULL,
      [mutation_documentation] [text] NULL,
      [mutation_comment] [text] NULL,
      [first_author] [varchar](24) NULL,
      [year] [varchar](5) NULL,
      [journal] [varchar](250) NULL,
      [mutation_documentation_type] [varchar](18) NULL,
      [link] [varchar](1024) NULL,
      [phenotype] [varchar](250) NULL,
      [zygosity] [varchar](24) NULL,
      [reference_id] [int] NULL,
      [countryiso1] [varchar](250) NULL,
      [countryiso2] [varchar](250) NULL,
      [countryiso3] [varchar](250) NULL,
      [countryiso4] [varchar](250) NULL,
      [countryiso5] [varchar](250) NULL,
      [countryiso6] [varchar](250) NULL,
```

```
	[countryiso7] [varchar](250) NULL,
	[countryiso8] [varchar](250) NULL,
	[countryiso9] [varchar](250) NULL,
	[countryiso10] [varchar](250) NULL,
	[countryiso11] [varchar](250) NULL,
	[countryiso12] [varchar](250) NULL,
	[countryiso13] [varchar](250) NULL,
	[countryiso14] [varchar](250) NULL,
	[countryiso15] [varchar](250) NULL,
	[countryiso16] [varchar](250) NULL,
	[countryiso17] [varchar](250) NULL,
	[countryiso18] [varchar](250) NULL,
	[countryiso19] [varchar](250) NULL,
	[countryiso20] [varchar](250) NULL,
	[swissprot_id] [varchar](50) NULL,
	[pubmed_id] [varchar](20) NULL,
	[title] [varchar](1024) NULL,
	[doi] [varchar](150) NULL,
	[HGVS_Nucleotide] [varchar](250) NULL,
	[HGVS_protein] [varchar](250) NULL,
	[Location][varchar](250) NULL,
	[dbSNP] [varchar](20) NULL,
	[frequency] [int] NULL,
	[frequency_references] [varchar](250) NULL,
	[data_from] [varchar](30) NULL)
GO

BULK INSERT dbo.[mutation_new]
FROM '\\vm3\Alsod\DataSets\mutation_new.txt'
WITH
(
FIELDTERMINATOR = '\t',
ROWTERMINATOR = '\n'
)
GO

---Insert HGVS_Nucleotide,HGVS_protein, Location columns
ALTER TABLE dbo.[mutation]
ADD [HGVS_Nucleotide] [varchar](250) NULL;
GO

ALTER TABLE dbo.[mutation]
ADD [HGVS_protein] [varchar](250) NULL;
GO

ALTER TABLE dbo.[mutation]
ADD [Location] [varchar](250) NULL;
GO

ALTER TABLE dbo.[mutation]
ADD [dbSNP_id] [varchar](250) NULL;
GO

---Update mutation table with data
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.33C>T',
HGVS_protein = 'NP_000445.1:p.G11G', Location ='21:33032115', dbSNP_id=''
WHERE mutation_id = '1'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.180T>C',
HGVS_protein = 'NP_000445.1:p.S60S', Location ='21:33038772', dbSNP_id=''
WHERE mutation_id = '2'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.423T>A',
HGVS_protein    =    'NP_000445.1:p.A141A',    Location    ='21:33040849',
dbSNP_id='rs143100660'  WHERE mutation_id = '3'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.420C>T',
HGVS_protein    =    'NP_000445.1:p.N140N',    Location    ='21:33040846',
dbSNP_id='rs1804449'  WHERE mutation_id = '4'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.462A>G',
HGVS_protein = 'NP_000445.1:p.Q154Q', Location ='21:33040888', dbSNP_id=''
WHERE mutation_id = '5'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '6'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '7'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.344G>C',
HGVS_protein = 'NP_000445.1:p.G115A', Location ='21:33039675', dbSNP_id=''
WHERE mutation_id = '8'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.269C>T',
HGVS_protein = 'NP_000445.1:p.A90V', Location ='21:33039600', dbSNP_id=''
WHERE mutation_id = '9'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.317C>T',
HGVS_protein = 'NP_000445.1:p.S106L', Location ='21:33039648', dbSNP_id=''
WHERE mutation_id = '10'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.340A>T',
HGVS_protein = 'NP_000445.1:p.I114F', Location ='21:33039671', dbSNP_id=''
WHERE mutation_id = '11'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.422C>G',
HGVS_protein = 'NP_000445.1:p.A141G', Location ='21:33040848', dbSNP_id=''
WHERE mutation_id = '12'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.137T>G',
HGVS_protein    =    'NP_000445.1:p.F46C',    Location    ='21:33036167',
dbSNP_id='rs121912457'  WHERE mutation_id = '13'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.286G>A',
HGVS_protein = 'NP_000445.1:p.A96T', Location ='21:33039617', dbSNP_id=''
WHERE mutation_id = '14'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.25C>G',
HGVS_protein = 'NP_000445.1:p.L9V', Location ='21:33032107', dbSNP_id=''
WHERE mutation_id = '15'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.261T>A',
HGVS_protein = 'NP_000445.1:p.N87K', Location ='21:33039592', dbSNP_id=''
WHERE mutation_id = '16'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.142G>T',
HGVS_protein = 'NP_000445.1:p.V48F', Location ='21:33036172', dbSNP_id=''
WHERE mutation_id = '17'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.230A>T',
HGVS_protein = 'NP_000445.1:p.D77V', Location ='21:33038822', dbSNP_id=''
WHERE mutation_id = '18'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '19'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.116T>G',
HGVS_protein = 'NP_000445.1:p.L39R', Location ='21:33036146', dbSNP_id=''
WHERE mutation_id = '20'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.115C>G',
HGVS_protein     =     'NP_000445.1:p.L39V',     Location     ='21:33036145',
dbSNP_id='rs121912432'  WHERE mutation_id = '21'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.124G>A',
HGVS_protein     =     'NP_000445.1:p.G42S',     Location     ='21:33036154',
dbSNP_id='rs121912433'  WHERE mutation_id = '22'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.125G>A',
HGVS_protein     =     'NP_000445.1:p.G42D',     Location     ='21:33036155',
dbSNP_id='rs121912434'  WHERE mutation_id = '24'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.131A>G',
HGVS_protein     =     'NP_000445.1:p.H44R',     Location     ='21:33036161',
dbSNP_id='rs121912435'  WHERE mutation_id = '25'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.140A>G',
HGVS_protein     =     'NP_000445.1:p.H47R',     Location     ='21:33036170',
dbSNP_id='rs121912443'  WHERE mutation_id = '26'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.148G>A',
HGVS_protein = 'NP_000445.1:p.E50K', Location ='21:33036178', dbSNP_id=''
WHERE mutation_id = '27'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.203T>G',
HGVS_protein = 'NP_000445.1:p.L68R', Location ='21:33038795', dbSNP_id=''
WHERE mutation_id = '28'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '29'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.255G>C',
HGVS_protein = 'NP_000445.1:p.L85F', Location ='21:33039586', dbSNP_id=''
WHERE mutation_id = '30'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.259A>G',
HGVS_protein = 'NP_000445.1:p.N87D', Location ='21:33039590', dbSNP_id=''
WHERE mutation_id = '31'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.256G>C',
HGVS_protein     =     'NP_000445.1:p.G86R',     Location     ='21:33039587',
dbSNP_id='rs121912436'  WHERE mutation_id = '32'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.260A>G',
HGVS_protein     =     'NP_000445.1:p.N87S',     Location     ='21:33039591',
dbSNP_id='rs11556620'  WHERE mutation_id = '33'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.272A>T',
HGVS_protein     =     'NP_000445.1:p.D91V',     Location     ='21:33039603',
dbSNP_id='rs80265967'  WHERE mutation_id = '34'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.281G>C',
HGVS_protein     =     'NP_000445.1:p.G94A',     Location     ='21:33039612',
dbSNP_id='rs121912438'  WHERE mutation_id = '35'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.280G>T',
HGVS_protein     =     'NP_000445.1:p.G94C',     Location     ='21:33039611',
dbSNP_id='rs121912437; rs121912460'  WHERE mutation_id = '36'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.280G>C',
HGVS_protein     =     'NP_000445.1:p.G94R',     Location     ='21:33039611',
dbSNP_id='rs121912437; rs121912460'  WHERE mutation_id = '37'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.281G>A',
HGVS_protein     =     'NP_000445.1:p.G94D',     Location     ='21:33039612',
dbSNP_id='rs121912438'  WHERE mutation_id = '38'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '39'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.304G>C',
HGVS_protein = 'NP_000445.1:p.D102H', Location ='21:33039635', dbSNP_id=''
WHERE mutation_id = '40'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.268G>A',
HGVS_protein = 'NP_000445.1:p.A90T', Location ='21:33039599', dbSNP_id=''
WHERE mutation_id = '41'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.418A>C',
HGVS_protein = 'NP_000445.1:p.N140H', Location ='21:33040844', dbSNP_id=''
WHERE mutation_id = '42'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.164C>G',
HGVS_protein = 'NP_000445.1:p.T55R', Location ='21:33036194', dbSNP_id=''
WHERE mutation_id = '43'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.147T>G',
HGVS_protein = 'NP_000445.1:p.H49Q', Location ='21:33036177', dbSNP_id=''
WHERE mutation_id = '44'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.376G>C',
HGVS_protein = 'NP_000445.1:p.D126H', Location ='21:33040802', dbSNP_id=''
WHERE mutation_id = '45'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.272A>C',
HGVS_protein     =     'NP_000445.1:p.D91A',     Location     ='21:33039603',
dbSNP_id='rs80265967'  WHERE mutation_id = '46'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.217G>A',
HGVS_protein     =     'NP_000445.1:p.G73S',     Location     ='21:33038809',
dbSNP_id='rs121912455'  WHERE mutation_id = '47'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.263T>C',
HGVS_protein = 'NP_000445.1:p.V88A', Location ='21:33039594', dbSNP_id=''
WHERE mutation_id = '48'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.43G>A',
HGVS_protein = 'NP_000445.1:p.V15M', Location ='21:33032125', dbSNP_id=''
WHERE mutation_id = '49'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.253T>G',
HGVS_protein     =     'NP_000445.1:p.L85V',     Location     ='21:33039584',
dbSNP_id='rs121912452'  WHERE mutation_id = '50'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.44T>G',
HGVS_protein = 'NP_000445.1:p.V15G', Location ='21:33032126', dbSNP_id=''
WHERE mutation_id = '51'
```

```
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.229G>T',
HGVS_protein  =  'NP_000445.1:p.D77Y',  Location  ='21:33038821',  dbSNP_id=''
WHERE mutation_id = '52'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '53'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '54'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '55'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.374A>T',
HGVS_protein  =  'NP_000445.1:p.D125V',  Location  ='21:33040800',  dbSNP_id=''
WHERE mutation_id = '56'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.179G>T',
HGVS_protein  =  'NP_000445.1:p.S60I',  Location  ='21:33038771',  dbSNP_id=''
WHERE mutation_id = '57'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.281G>T',
HGVS_protein     =     'NP_000445.1:p.G94V',     Location     ='21:33039612',
dbSNP_id='rs121912438'  WHERE mutation_id = '58'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.425G>A',
HGVS_protein  =  'NP_000445.1:p.G142E',  Location  ='21:33040851',  dbSNP_id=''
WHERE mutation_id = '59'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.420C>A',
HGVS_protein     =     'NP_000445.1:p.N140K',     Location     ='21:33040846',
dbSNP_id='rs1804449'  WHERE mutation_id = '60'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.449T>C',
HGVS_protein  =  'NP_000445.1:p.I150T',  Location  ='21:33040875',  dbSNP_id=''
WHERE mutation_id = '61'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.455T>C',
HGVS_protein     =     'NP_000445.1:p.I152T',     Location     ='21:33040881',
dbSNP_id='rs121912449'  WHERE mutation_id = '62'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '63'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.19T>G',
HGVS_protein  =  'NP_000445.1:p.C7G',  Location  ='21:33032101',  dbSNP_id=''
WHERE mutation_id = '64'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '65'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.455T>G',
HGVS_protein     =     'NP_000445.1:p.I152S',     Location     ='21:33040881',
dbSNP_id='rs121912449'  WHERE mutation_id = '66'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.13G>A',
HGVS_protein     =     'NP_000445.1:p.A5T',     Location     ='21:33032095',
dbSNP_id='rs121912444'  WHERE mutation_id = '67'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '68'
```

```
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.355G>C',
HGVS_protein = 'NP_000445.1:p.V119L', Location ='21:33039686', dbSNP_id=''
WHERE mutation_id = '69'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.62T>G',
HGVS_protein = 'NP_000445.1:p.F21C', Location ='21:33032144', dbSNP_id=''
WHERE mutation_id = '70'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.424G>T',
HGVS_protein = 'NP_000445.1:p.G142X', Location ='21:33040850', dbSNP_id=''
WHERE mutation_id = '71'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '72'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.59A>G',
HGVS_protein = 'NP_000445.1:p.N20S', Location ='21:33032141', dbSNP_id=''
WHERE mutation_id = '73'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '74'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.442G>C',
HGVS_protein = 'NP_000445.1:p.G148R', Location ='21:33040868', dbSNP_id=''
WHERE mutation_id = '75'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.242A>G',
HGVS_protein    =    'NP_000445.1:p.H81R',    Location    ='21:33039573',
dbSNP_id='rs121912458'  WHERE mutation_id = '76'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '77'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.304G>A',
HGVS_protein = 'NP_000445.1:p.D102N', Location ='21:33039635', dbSNP_id=''
WHERE mutation_id = '78'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.313A>T',
HGVS_protein    =    'NP_000445.1:p.I105F',    Location    ='21:33039644',
dbSNP_id='rs121912445'  WHERE mutation_id = '79'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.319C>G',
HGVS_protein    =    'NP_000445.1:p.L107V',    Location    ='21:33039650',
dbSNP_id='rs121912440'  WHERE mutation_id = '80'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.338T>C',
HGVS_protein    =    'NP_000445.1:p.I113T',    Location    ='21:33039669',
dbSNP_id='rs74315452'  WHERE mutation_id = '81'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.341T>C',
HGVS_protein    =    'NP_000445.1:p.I114T',    Location    ='21:33039672',
dbSNP_id='rs121912441'  WHERE mutation_id = '82'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.404G>A',
HGVS_protein    =    'NP_000445.1:p.S135N',    Location    ='21:33040830',
dbSNP_id='rs121912451'  WHERE mutation_id = '83'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.435G>C',
HGVS_protein = 'NP_000445.1:p.L145F', Location ='21:33040861', dbSNP_id=''
WHERE mutation_id = '84'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.434T>C',
HGVS_protein    =    'NP_000445.1:p.L145S',    Location    ='21:33040860',
dbSNP_id='rs121912446'  WHERE mutation_id = '85'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.436G>A',
HGVS_protein    =    'NP_000445.1:p.A146T',    Location    ='21:33040862',
dbSNP_id='rs121912447'  WHERE mutation_id = '86'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.439T>C',
HGVS_protein = 'NP_000445.1:p.C147R', Location ='21:33040865', dbSNP_id=''
WHERE mutation_id = '87'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.446T>G',
HGVS_protein = 'NP_000445.1:p.V149G', Location ='21:33040872', dbSNP_id=''
WHERE mutation_id = '88'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.197A>G',
HGVS_protein = 'NP_000445.1:p.N66S', Location ='21:33038789', dbSNP_id=''
WHERE mutation_id = '89'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.339C>G',
HGVS_protein = 'NP_000445.1:p.I113M', Location ='21:33039670', dbSNP_id=''
WHERE mutation_id = '90'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.289G>A',
HGVS_protein    =    'NP_000445.1:p.D97N',    Location    ='21:33039620',
dbSNP_id='rs121912459'  WHERE mutation_id = '91'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '92'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.445G>A',
HGVS_protein = 'NP_000445.1:p.V149I', Location ='21:33040871', dbSNP_id=''
WHERE mutation_id = '93'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '94'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.380T>A',
HGVS_protein    =    'NP_000445.1:p.L127X',    Location    ='21:33040806',
dbSNP_id='rs121912454'  WHERE mutation_id = '95'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.374A>T',
HGVS_protein = 'NP_000445.1:p.D125V', Location ='21:33040800', dbSNP_id=''
WHERE mutation_id = '96'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.326G>T',
HGVS_protein = 'NP_000445.1:p.G109V', Location ='21:33039657', dbSNP_id=''
WHERE mutation_id = '97'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004738.4:c.166C>T',
HGVS_protein    =    'NP_004729.1:p.P56S',    Location    ='20:56993374',
dbSNP_id='rs74315431'  WHERE mutation_id = '98'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.346C>G',
HGVS_protein    =    'NP_000445.1:p.R116G',    Location    ='21:33039677',
dbSNP_id='rs74315431'  WHERE mutation_id = '99'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '100'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.14C>T',
HGVS_protein     =     'NP_000445.1:p.A5V',     Location     ='21:33032096',
dbSNP_id='rs121912442'  WHERE mutation_id = '101'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.13G>T',
HGVS_protein     =     'NP_000445.1:p.A5S',     Location     ='21:33032095',
dbSNP_id='rs121912444'  WHERE mutation_id = '102'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '103'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.23T>A',
HGVS_protein  =  'NP_000445.1:p.V8E',  Location  ='21:33032105',  dbSNP_id=''
WHERE mutation_id = '104'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.26T>A',
HGVS_protein  =  'NP_000445.1:p.L9Q',  Location  ='21:33032108',  dbSNP_id=''
WHERE mutation_id = '105'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.37G>C',
HGVS_protein     =     'NP_000445.1:p.G13R',     Location     ='21:33032119',
dbSNP_id='rs121912456'  WHERE mutation_id = '106'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.50G>C',
HGVS_protein  =  'NP_000445.1:p.G17A',  Location  ='21:33032132',  dbSNP_id=''
WHERE mutation_id = '107'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.49G>A',
HGVS_protein     =     'NP_000445.1:p.G17S',     Location     ='21:33032131',
dbSNP_id='rs121912453'  WHERE mutation_id = '108'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.64G>A',
HGVS_protein     =     'NP_000445.1:p.E22K',     Location     ='21:33032146',
dbSNP_id='rs121912450'  WHERE mutation_id = '109'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.65A>G',
HGVS_protein  =  'NP_000445.1:p.E22G',  Location  ='21:33032147',  dbSNP_id=''
WHERE mutation_id = '110'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.112G>A',
HGVS_protein     =     'NP_000445.1:p.G38R',     Location     ='21:33036142',
dbSNP_id='rs121912431'  WHERE mutation_id = '111'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.280G>A',
HGVS_protein     =     'NP_000445.1:p.G94S',     Location     ='21:33039611',
dbSNP_id='rs121912437; rs121912460'  WHERE mutation_id = '112'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.292G>A',
HGVS_protein  =  'NP_000445.1:p.V98M',  Location  ='21:33039623',  dbSNP_id=''
WHERE mutation_id = '113'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.302A>G',
HGVS_protein     =     'NP_000445.1:p.E101G',     Location     ='21:33039633',
dbSNP_id='rs121912439'  WHERE mutation_id = '114'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.301G>A',
HGVS_protein  =  'NP_000445.1:p.E101K',  Location  ='21:33039632',  dbSNP_id=''
WHERE mutation_id = '115'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.305A>G',
HGVS_protein  =  'NP_000445.1:p.D102G',  Location  ='21:33039636',  dbSNP_id=''
WHERE mutation_id = '116'
```

```
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.217G>T',
HGVS_protein     =     'NP_000445.1:p.G73C',     Location     ='21:33038809',
dbSNP_id='rs121912455'  WHERE mutation_id = '117'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '118'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.437C>G',
HGVS_protein = 'NP_000445.1:p.A146G', Location ='21:33040863', dbSNP_id=''
WHERE mutation_id = '119'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.68A>T',
HGVS_protein = 'NP_000445.1:p.Q23L', Location ='21:33032150', dbSNP_id=''
WHERE mutation_id = '120'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.146A>G',
HGVS_protein = 'NP_000445.1:p.H49R', Location ='21:33036176', dbSNP_id=''
WHERE mutation_id = '121'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.352C>G',
HGVS_protein     =     'NP_000445.1:p.L118V',     Location     ='21:33039683',
dbSNP_id='rs199474723'  WHERE mutation_id = '122'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.31G>C',
HGVS_protein = 'NP_000445.1:p.G11R', Location ='21:33032113', dbSNP_id=''
WHERE mutation_id = '123'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.298A>G',
HGVS_protein = 'NP_000445.1:p.I100V', Location ='21:33039629', dbSNP_id=''
WHERE mutation_id = '124'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.290A>T',
HGVS_protein = 'NP_000445.1:p.D97V', Location ='21:33039621', dbSNP_id=''
WHERE mutation_id = '125'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.287C>T',
HGVS_protein = 'NP_000445.1:p.A96V', Location ='21:33039618', dbSNP_id=''
WHERE mutation_id = '126'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.304G>T',
HGVS_protein = 'NP_000445.1:p.D102Y', Location ='21:33039635', dbSNP_id=''
WHERE mutation_id = '127'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.355G>T',
HGVS_protein = 'NP_000445.1:p.V119L', Location ='21:33039686', dbSNP_id=''
WHERE mutation_id = '129'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.380T>C',
HGVS_protein     =     'NP_000445.1:p.L127S',     Location     ='21:33040806',
dbSNP_id='rs121912454'  WHERE mutation_id = '130'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '131'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '134'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '135'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '136'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.89T>C',
HGVS_protein = 'NP_000445.1:p.V30A', Location ='21:33036119', dbSNP_id=''
WHERE mutation_id = '137'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.143T>C',
HGVS_protein = 'NP_000445.1:p.V48A', Location ='21:33036173', dbSNP_id=''
WHERE mutation_id = '138'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.401A>T',
HGVS_protein = 'NP_000445.1:p.E134V', Location ='21:33040827', dbSNP_id=''
WHERE mutation_id = '139'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.260A>T',
HGVS_protein    =    'NP_000445.1:p.N87I',    Location    ='21:33039591',
dbSNP_id='rs11556620'  WHERE mutation_id = '140'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.16G>T',
HGVS_protein = 'NP_000445.1:p.V6L', Location ='21:33032098', dbSNP_id=''
WHERE mutation_id = '141'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.199C>G',
HGVS_protein = 'NP_000445.1:p.P67A', Location ='21:33038791', dbSNP_id=''
WHERE mutation_id = '142'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.262G>A',
HGVS_protein = 'NP_000445.1:p.V88M', Location ='21:33039593', dbSNP_id=''
WHERE mutation_id = '145'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.292G>C',
HGVS_protein = 'NP_000445.1:p.V98L', Location ='21:33039623', dbSNP_id=''
WHERE mutation_id = '146'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.418A>G',
HGVS_protein = 'NP_000445.1:p.N140D', Location ='21:33040844', dbSNP_id=''
WHERE mutation_id = '147'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.443G>A',
HGVS_protein = 'NP_000445.1:p.G148D', Location ='21:33040869', dbSNP_id=''
WHERE mutation_id = '148'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.351A>G',
HGVS_protein = 'NP_000445.1:p.T117T', Location ='21:33039682', dbSNP_id=''
WHERE mutation_id = '149'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.441T>A',
HGVS_protein = 'NP_000445.1:p.C147X', Location ='21:33040867', dbSNP_id=''
WHERE mutation_id = '150'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.256G>A',
HGVS_protein    =    'NP_000445.1:p.G86S',    Location    ='21:33039587',
dbSNP_id='rs121912436'  WHERE mutation_id = '151'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.107A>T',
HGVS_protein    =    'NP_001136.1:p.Q36L',    Location    ='14:21161830',
dbSNP_id='rs121909535'  WHERE mutation_id = '153'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.122A>T',
HGVS_protein    =    'NP_001136.1:p.K41I',    Location    ='14:21161845',
dbSNP_id='rs121909536'  WHERE mutation_id = '154'
```

```
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.250A>G ',
HGVS_protein = ' NP_001136.1:p.Lys84Glu', Location ='', dbSNP_id='rs17560'
WHERE mutation_id = '155'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.164G>A',
HGVS_protein     =     'NP_001136.1:p.R55K',     Location     ='14:21161887',
dbSNP_id='rs121909538'  WHERE mutation_id = '157'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='',    HGVS_protein    =
' NP_001136.1:p.Cys39Trp', Location ='', dbSNP_id=''   WHERE mutation_id =
'160'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.191A>T',
HGVS_protein     =     'NP_001136.1:p.K64I',     Location     ='14:21161914',
dbSNP_id='rs121909540'  WHERE mutation_id = '161'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='',    HGVS_protein    =
' NP_001136.1:p.Ile46Val', Location ='', dbSNP_id=''   WHERE mutation_id =
'162'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.155G>A',
HGVS_protein     =     'NP_001136.1:p.S52N',     Location     ='14:21161878',
dbSNP_id='rs121909542'  WHERE mutation_id = '164'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.407C>T',
HGVS_protein     =     'NP_001136.1:p.P136L',     Location     ='14:21162130',
dbSNP_id='rs121909543'  WHERE mutation_id = '165'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.409G>A',
HGVS_protein     =     'NP_001136.1:p.V137I',     Location     ='14:21162132',
dbSNP_id='rs121909544'  WHERE mutation_id = '166'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1009A>G',
HGVS_protein     =     'NP_031401.1:p.M337V',     Location     ='1:11082475',
dbSNP_id='rs80356730'  WHERE mutation_id = '167'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.991C>A',
HGVS_protein     =     'NP_031401.1:p.Q331K',     Location     ='1:11082457',
dbSNP_id='rs80356727'  WHERE mutation_id = '168'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.881G>C',
HGVS_protein     =     'NP_031401.1:p.G294A',     Location     ='1:11082347',
dbSNP_id='rs80356721'  WHERE mutation_id = '169'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.869G>C',
HGVS_protein     =     'NP_031401.1:p.G290A',     Location     ='1:11082335',
dbSNP_id='rs121908395'  WHERE mutation_id = '170'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.892G>A',
HGVS_protein     =     'NP_031401.1:p.G298S',     Location     ='1:11082358',
dbSNP_id='rs4884357'  WHERE mutation_id = '171'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.506A>G',
HGVS_protein     =     'NP_031401.1:p.D169G',     Location     ='1:11078893',
dbSNP_id='rs80356717'  WHERE mutation_id = '172'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1042G>T',
HGVS_protein     =     'NP_031401.1:p.G348C',     Location     ='1:11082508',
dbSNP_id='rs80356733'  WHERE mutation_id = '173'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1028A>G',
HGVS_protein     =     'NP_031401.1:p.Q343R',     Location     ='1:11082494',
dbSNP_id='rs80356731'  WHERE mutation_id = '174'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.943G>A',
HGVS_protein     =     'NP_031401.1:p.A315T',     Location     ='1:11082409',
dbSNP_id='rs80356726'  WHERE mutation_id = '175'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1562G>A',
HGVS_protein     =     'NP_004951.1:p.R521H',     Location     ='16:31202740',
dbSNP_id='rs121909671'  WHERE mutation_id = '176'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1561C>T',
HGVS_protein     =     'NP_004951.1:p.R521C',     Location     ='16:31202739',
dbSNP_id='rs121909668; rs121909670'  WHERE mutation_id = '177'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1540A>G',
HGVS_protein = 'NP_004951.1:p.R514G', Location ='16:31202430', dbSNP_id=''
WHERE mutation_id = '178'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1551C>G',
HGVS_protein     =     'NP_004951.1:p.H517Q',     Location     ='16:31202729',
dbSNP_id='rs121909667'  WHERE mutation_id = '179'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '180'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.730C>T',
HGVS_protein = 'NP_004951.1:p.R244C', Location ='16:31196466', dbSNP_id=''
WHERE mutation_id = '181'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1542G>T',
HGVS_protein = 'NP_004951.1:p.R514S', Location ='16:31202720', dbSNP_id=''
WHERE mutation_id = '182'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1543G>T',
HGVS_protein = 'NP_004951.1:p.G515C', Location ='16:31202721', dbSNP_id=''
WHERE mutation_id = '183'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1553G>A',
HGVS_protein     =     'NP_004951.1:p.R518K',     Location     ='16:31202731',
dbSNP_id='rs121909669'  WHERE mutation_id = '184'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1561C>G',
HGVS_protein     =     'NP_004951.1:p.R521G',     Location     ='16:31202739',
dbSNP_id='rs121909668; rs121909670'  WHERE mutation_id = '185'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1564A>G',
HGVS_protein = 'NP_004951.1:p.R522G', Location ='16:31202742', dbSNP_id=''
WHERE mutation_id = '186'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1571G>C',
HGVS_protein = 'NP_004951.1:p.R524T', Location ='16:31202749', dbSNP_id=''
WHERE mutation_id = '187'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1572G>C',
HGVS_protein = 'NP_004951.1:p.R524S', Location ='16:31202750', dbSNP_id=''
WHERE mutation_id = '188'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1574C>T',
HGVS_protein = 'NP_004951.1:p.P525L', Location ='16:31202752', dbSNP_id=''
WHERE mutation_id = '189'
```

```
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '190'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '195'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '196'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '197'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '198'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '199'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '200'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '201'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.2992C>T',
HGVS_protein    =    'NP_065970.2:p.R998X',    Location    ='2:202591577',
dbSNP_id='rs121908137'  WHERE mutation_id = '202'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_001145.4:c.3G>A', HGVS_protein
=  'NP_001136.1:p.M1I',  Location  ='14:21161726',  dbSNP_id='rs201068740'
WHERE mutation_id = '204'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.35T>C',
HGVS_protein = 'NP_001136.1:p.F12S', Location ='14:21161758', dbSNP_id=''
WHERE mutation_id = '205'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.61C>T',
HGVS_protein    =    'NP_001136.1:p.P21S',    Location    ='14:21161784',
dbSNP_id='rs149672657'  WHERE mutation_id = '206'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.413A>G',
HGVS_protein = 'NP_001136.1:p.H138R', Location ='14:21162136', dbSNP_id=''
WHERE mutation_id = '207'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.132C>G',
HGVS_protein    =    'NP_001136.1:p.G44G',    Location    ='14:21161855',
dbSNP_id='rs147701713'  WHERE mutation_id = '208'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.883G>A',
HGVS_protein    =    'NP_031401.1:p.G295S',    Location    ='1:11082349',
dbSNP_id='rs80356723'  WHERE mutation_id = '209'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.269C>T',
HGVS_protein    =    'NP_031401.1:p.A90V',    Location    ='1:11076931',
dbSNP_id='rs80356715'  WHERE mutation_id = '210'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.800A>G',
HGVS_protein    =    'NP_031401.1:p.N267S',    Location    ='1:11082266',
dbSNP_id='rs80356718'  WHERE mutation_id = '211'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.859G>A',
HGVS_protein     =     'NP_031401.1:p.G287S',     Location     ='1:11082325',
dbSNP_id='rs80356719'  WHERE mutation_id = '212'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1083G>T',
HGVS_protein     =     'NP_031401.1:p.R361S',     Location     ='1:11082549',
dbSNP_id='rs80356735'  WHERE mutation_id = '213'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1144G>A',
HGVS_protein     =     'NP_031401.1:p.A382T',     Location     ='1:11082610',
dbSNP_id='rs11689432'  WHERE mutation_id = '214'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1168A>G',
HGVS_protein     =     'NP_031401.1:p.N390D',     Location     ='1:11082634',
dbSNP_id='rs80356741'  WHERE mutation_id = '215'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1169A>G',
HGVS_protein     =     'NP_031401.1:p.N390S',     Location     ='1:11082635',
dbSNP_id='rs80356742'  WHERE mutation_id = '216'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.881G>T',
HGVS_protein     =     'NP_031401.1:p.G294V',     Location     ='1:11082347',
dbSNP_id='rs80356721'  WHERE mutation_id = '217'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.883G>C',
HGVS_protein     =     'NP_031401.1:p.G295R',     Location     ='1:11082349',
dbSNP_id='rs80356723'  WHERE mutation_id = '218'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.995G>A',
HGVS_protein     =     'NP_031401.1:p.S332N',     Location     ='1:11082461',
dbSNP_id='rs80356728'  WHERE mutation_id = '219'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1004G>A',
HGVS_protein     =     'NP_031401.1:p.G335D',     Location     ='1:11082470',
dbSNP_id='rs80356729'  WHERE mutation_id = '220'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1135T>C',
HGVS_protein     =     'NP_031401.1:p.S379P',     Location     ='1:11082601',
dbSNP_id='rs80356738'  WHERE mutation_id = '221'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1136C>G',
HGVS_protein     =     'NP_031401.1:p.S379C',     Location     ='1:11082602',
dbSNP_id='rs80356739'  WHERE mutation_id = '222'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1178C>T',
HGVS_protein     =     'NP_031401.1:p.S393L',     Location     ='1:11082644',
dbSNP_id='rs80356743'  WHERE mutation_id = '223'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.319C>T',
HGVS_protein     =     'NP_000445.1:p.L107F',     Location     ='21:33039650',
dbSNP_id='rs121912440'  WHERE mutation_id = '224'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.280A>G',
HGVS_protein     =     'NP_065970.2:p.I94V',     Location     ='2:202626437',
dbSNP_id='rs3219154'  WHERE mutation_id = '225'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.305A>G',
HGVS_protein = 'NP_065970.2:p.H102R', Location ='2:202626412', dbSNP_id=''
WHERE mutation_id = '226'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_015046.5:c.1166T>C',
HGVS_protein = 'NP_055861.3:p.L389S', Location ='9:135205819',
dbSNP_id='rs29001584' WHERE mutation_id = '227'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_015046.5:c.6407G>A',
HGVS_protein = 'NP_055861.3:p.R2136H', Location ='9:135158790',
dbSNP_id='rs121434378' WHERE mutation_id = '228'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_015046.5:c.8C>T', HGVS_protein
= 'NP_055861.3:p.T3I', Location ='9:135224808', dbSNP_id='rs28941475'
WHERE mutation_id = '229'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_007375.3:c.1035C>A',
HGVS_protein = 'NP_031401.1:p.N345K', Location ='1:11082501',
dbSNP_id='rs80356732' WHERE mutation_id = '230'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_007375.3:c.1147A>G',
HGVS_protein = 'NP_031401.1:p.I383V', Location ='1:11082613',
dbSNP_id='rs80356740' WHERE mutation_id = '231'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_007375.3:c.1055A>G',
HGVS_protein = 'NP_031401.1:p.N352S', Location ='1:11082521',
dbSNP_id='rs80356734' WHERE mutation_id = '232'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_007375.3:c.1087C>G',
HGVS_protein = 'NP_031401.1:p.P363A', Location ='1:11082553', dbSNP_id=''
WHERE mutation_id = '233'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_007375.3:c.1122T>A',
HGVS_protein = 'NP_031401.1:p.Y374X', Location ='1:11082588',
dbSNP_id='rs147795017' WHERE mutation_id = '234'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_007375.3:c.1144G>C',
HGVS_protein = 'NP_031401.1:p.A382P', Location ='1:11082610',
dbSNP_id='rs11689432' WHERE mutation_id = '235'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.169_171delTCT',
HGVS_protein = 'NP_004951.1:p.S57del', Location ='16:31101465', dbSNP_id=''
WHERE mutation_id = '236'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.467G>A',
HGVS_protein = 'NP_004951.1:p.G156E', Location ='16:31195661', dbSNP_id=''
WHERE mutation_id = '237'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.701G>T',
HGVS_protein = 'NP_004951.1:p.R234L', Location ='16:31196437', dbSNP_id=''
WHERE mutation_id = '238'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_007375.3:c.931A>G',
HGVS_protein = 'NP_031401.1:p.M311V', Location ='1:11082397',
dbSNP_id='rs80356725' WHERE mutation_id = '239'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_000454.4:c.382G>C',
HGVS_protein = 'NP_000445.1:p.G128R', Location ='21:33040808', dbSNP_id=''
WHERE mutation_id = '240'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_007375.3:c.962C>G',
HGVS_protein = 'NP_031401.1:p.A321G', Location ='1:11082428', dbSNP_id=''
WHERE mutation_id = '241'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1520G>A',
HGVS_protein = 'NP_004951.1:p.G507D', Location ='16:31202410', dbSNP_id=''
WHERE mutation_id = '242'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.413C>G',
HGVS_protein = 'NP_000445.1:p.T138R', Location ='21:33040839', dbSNP_id=''
WHERE mutation_id = '243'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.416G>A',
HGVS_protein = 'NP_000445.1:p.G139E', Location ='21:33040842', dbSNP_id=''
WHERE mutation_id = '244'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1549C>G',
HGVS_protein = 'NP_004951.1:p.H517D', Location ='16:31202727', dbSNP_id=''
WHERE mutation_id = '245'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1132A>G',
HGVS_protein = 'NP_031401.1:p.N378D', Location ='1:11082598', dbSNP_id=''
WHERE mutation_id = '246'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.700C>T',
HGVS_protein = 'NP_004951.1:p.R234C', Location ='16:31196436', dbSNP_id=''
WHERE mutation_id = '247'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.688G>T',
HGVS_protein = 'NP_004951.1:p.G230C', Location ='16:31196424', dbSNP_id=''
WHERE mutation_id = '248'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.674G>T',
HGVS_protein = 'NP_004951.1:p.G225V', Location ='16:31196410', dbSNP_id=''
WHERE mutation_id = '249'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.646C>T',
HGVS_protein = 'NP_004951.1:p.R216C', Location ='16:31196382', dbSNP_id=''
WHERE mutation_id = '250'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.571G>A',
HGVS_protein    =    'NP_004951.1:p.G191S',    Location    ='16:31196307',
dbSNP_id='rs148758737'  WHERE mutation_id = '251'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.875G>A',
HGVS_protein = 'NP_031401.1:p.S292N', Location ='1:11082341', dbSNP_id=''
WHERE mutation_id = '252'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.120G>A',
HGVS_protein = 'NP_031401.1:p.G40G', Location ='1:11073904', dbSNP_id=''
WHERE mutation_id = '253'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1098C>G',
HGVS_protein    =    'NP_031401.1:p.A366A',    Location    ='1:11082564',
dbSNP_id='rs148325203'  WHERE mutation_id = '254'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_015046.5:c.3353C>T',
HGVS_protein = 'NP_055861.3:p.T1118I', Location ='9:135203632', dbSNP_id=''
WHERE mutation_id = '255'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1561C>A',
HGVS_protein    =    'NP_004951.1:p.R521S',    Location    ='16:31202739',
dbSNP_id='rs121909668; rs121909670'  WHERE mutation_id = '256'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1562G>T',
HGVS_protein     =     'NP_004951.1:p.R521L',     Location     ='16:31202740',
dbSNP_id='rs121909671'  WHERE mutation_id = '257'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.335G>A',
HGVS_protein = 'NP_000445.1:p.C112Y', Location ='21:33039666', dbSNP_id=''
WHERE mutation_id = '258'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.19T>A',
HGVS_protein = 'NP_000445.1:p.C7S', Location ='21:33032101', dbSNP_id=''
WHERE mutation_id = '259'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021980.4:c.1192C>T',
HGVS_protein = 'NP_068815.2:p.Q398X', Location ='10:13167989', dbSNP_id=''
WHERE mutation_id = '260'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021980.4:c.1433A>G',
HGVS_protein = 'NP_068815.2:p.E478G', Location ='10:13174098', dbSNP_id=''
WHERE mutation_id = '261'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.287C>G',
HGVS_protein = 'NP_000445.1:p.A96G', Location ='21:33039618', dbSNP_id=''
WHERE mutation_id = '263'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.68A>G',
HGVS_protein = 'NP_000445.1:p.Q23R', Location ='21:33032150', dbSNP_id=''
WHERE mutation_id = '264'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.199C>T',
HGVS_protein = 'NP_000445.1:p.P67S', Location ='21:33038791', dbSNP_id=''
WHERE mutation_id = '265'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.184G>C',
HGVS_protein = 'NP_000445.1:p.G62R', Location ='21:33038776', dbSNP_id=''
WHERE mutation_id = '266'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '267'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '268'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1540A>C',
HGVS_protein = 'NP_004951.1:p.R514R', Location ='16:31202430', dbSNP_id=''
WHERE mutation_id = '270'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.2143C>T',
HGVS_protein     =     'NP_065970.2:p.Q715X',     Location     ='2:202609008',
dbSNP_id='rs121908139'  WHERE mutation_id = '271'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.1102G>A',
HGVS_protein     =     'NP_065970.2:p.V368M',     Location     ='2:202625615',
dbSNP_id='rs3219156'  WHERE mutation_id = '272'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.4217G>A',
HGVS_protein = 'NP_065970.2:p.R1406K', Location ='2:202574667', dbSNP_id=''
WHERE mutation_id = '273'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.4015C>T',
HGVS_protein     =     'NP_065970.2:p.L1339L',     Location     ='2:202575821',
dbSNP_id='rs3219168'  WHERE mutation_id = '274'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '276'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1150G>C',
HGVS_protein = 'NP_031401.1:p.G384R', Location ='1:11082616', dbSNP_id=''
WHERE mutation_id = '284'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1153T>G',
HGVS_protein = 'NP_031401.1:p.W385G', Location ='1:11082619', dbSNP_id=''
WHERE mutation_id = '285'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.139C>G',
HGVS_protein = 'NP_000445.1:p.H47D', Location ='21:33036169', dbSNP_id=''
WHERE mutation_id = '286'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.200C>G',
HGVS_protein = 'NP_000445.1:p.P67R', Location ='21:33038792', dbSNP_id=''
WHERE mutation_id = '287'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.251A>G',
HGVS_protein = 'NP_000445.1:p.D84G', Location ='21:33039582', dbSNP_id=''
WHERE mutation_id = '288'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007126.3:c.572G>A',
HGVS_protein     =     'NP_009057.1:p.R191Q',     Location     ='9:35065252',
dbSNP_id='rs121909334'  WHERE mutation_id = '294'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007126.3:c.475C>G',
HGVS_protein = 'NP_009057.1:p.R159G', Location ='9:35065349', dbSNP_id=''
WHERE mutation_id = '295'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007126.3:c.1774G>A',
HGVS_protein = 'NP_009057.1:p.D592N', Location ='9:35059720', dbSNP_id=''
WHERE mutation_id = '296'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007126.3:c.464G>A',
HGVS_protein     =     'NP_009057.1:p.R155H',     Location     ='9:35065360',
dbSNP_id='rs121909329; rs121909333'  WHERE mutation_id = '299'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '300'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001917.4:c.595C>T',
HGVS_protein     =     'NP_001908.3:p.R199W',     Location     ='12:109288126',
dbSNP_id='rs3825251'  WHERE mutation_id = '301'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '302'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '303'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '304'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_014845.5:c.157G>T',
HGVS_protein     =     'NP_055660.1:p.D53Y',     Location     ='6:110036371',
dbSNP_id='rs121908290'  WHERE mutation_id = '306'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_014845.5:c.143A>G',
HGVS_protein = 'NP_055660.1:p.D48G', Location ='6:110036357', dbSNP_id=''
WHERE mutation_id = '307'
```

```
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_014845.5:c.1162A>G',
HGVS_protein = 'NP_055660.1:p.R388G', Location ='6:110081477', dbSNP_id=''
WHERE mutation_id = '308'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_014845.5:c.1231A>G',
HGVS_protein = 'NP_055660.1:p.I411V', Location ='6:110081546', dbSNP_id=''
WHERE mutation_id = '309'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '310'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_014845.5:c.2705T>C',
HGVS_protein = 'NP_055660.1:p.I902T', Location ='6:110146449', dbSNP_id=''
WHERE mutation_id = '311'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '312'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '313'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.232A>G',
HGVS_protein    =    'NP_001136.1:p.K78E',    Location    ='14:21161955',
dbSNP_id='rs141055235'  WHERE mutation_id = '314'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.36C>T',
HGVS_protein    =    'NP_001136.1:p.F(-13)L',    Location    ='14:21161759',
dbSNP_id=''  WHERE mutation_id = '315'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.434G>A',
HGVS_protein = 'NP_001136.1:p.R145H', Location ='14:21162157', dbSNP_id=''
WHERE mutation_id = '317'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.203T>C',
HGVS_protein = 'NP_000445.1:p.L68P', Location ='21:33038795', dbSNP_id=''
WHERE mutation_id = '318'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.2464G>A',
HGVS_protein    =    'NP_065970.2:p.V822M',    Location    ='2:202598115',
dbSNP_id='rs2276615'  WHERE mutation_id = '319'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_020919.3:c.1619G>A',
HGVS_protein = 'NP_065970.2:p.G540E', Location ='2:202619247', dbSNP_id=''
WHERE mutation_id = '320'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004738.4:c.137C>T',
HGVS_protein = 'NP_004729.1:p.T46I', Location ='20:56993345', dbSNP_id=''
WHERE mutation_id = '321'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '325'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_002345.3:c.596T>C',
HGVS_protein    =    'NP_002336.1:p.L199P',    Location    ='12:91502161',
dbSNP_id='rs147975710'  WHERE mutation_id = '326'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001888.3:c.505C>T',
HGVS_protein    =    'NP_001879.1:p.R169C',    Location    ='16:21279043',
dbSNP_id='rs148349291'  WHERE mutation_id = '327'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004082.4:c.3746C>T',
HGVS_protein     =     'NP_004073.2:p.T1249I',     Location     ='2:74588717',
dbSNP_id='rs72466496'  WHERE mutation_id = '328'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.198T>C',
HGVS_protein     =     'NP_031401.1:p.A66A',     Location     ='1:11073982',
dbSNP_id='rs61730366'  WHERE mutation_id = '329'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004082.4:c.1712T>C',
HGVS_protein     =     'NP_004073.2:p.M571T',     Location     ='2:74595997',
dbSNP_id='rs121909343'  WHERE mutation_id = '330'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004082.4:c.2353C>T',
HGVS_protein     =     'NP_004073.2:p.R785W',     Location     ='2:74594023',
dbSNP_id='rs121909344'  WHERE mutation_id = '331'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004082.4:c.3302G>A',
HGVS_protein     =     'NP_004073.2:p.R1101K',     Location     ='2:74590464',
dbSNP_id='rs121909345'  WHERE mutation_id = '332'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.34G>T',
HGVS_protein  =  'NP_000445.1:p.D12Y',  Location  ='21:33032116',  dbSNP_id=''
WHERE mutation_id = '333'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.328G>T',
HGVS_protein  =  'NP_000445.1:p.D110Y',  Location  ='21:33039659',  dbSNP_id=''
WHERE mutation_id = '334'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '335'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_015046.5:c.4660T>G',
HGVS_protein     =     'NP_055861.3:p.C1554G',     Location     ='9:135202325',
dbSNP_id='rs112089123'  WHERE mutation_id = '336'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_015046.5:c.7640T>C',
HGVS_protein     =     'NP_055861.3:p.I2547T',     Location     ='9:135140020',
dbSNP_id='rs151117904'  WHERE mutation_id = '337'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='NM_000454.4:c.13_14delinsTT',
HGVS_protein  =  'NP_000445.1:p.A4F',  Location  ='21:33032095  -  21:33032096',
dbSNP_id=''  WHERE mutation_id = '338'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_018008.3:c.563G>A',
HGVS_protein     =     'NP_060478.3:p.G188D',     Location     ='3:62357981',
dbSNP_id='rs199850439'  WHERE mutation_id = '339'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003585.3:c.626G>T',
HGVS_protein  =  'NP_003576.2:p.R209L',  Location  ='17:11884',  dbSNP_id=''
WHERE mutation_id = '340'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_014461.2:c.940T>G',
HGVS_protein     =     'NP_055276.1:p.F314V',     Location     ='3:1363512',
dbSNP_id='rs144474590'  WHERE mutation_id = '341'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_014461.2:c.2861A>T',
HGVS_protein  =  'NP_055276.1:p.E954V',  Location  ='3:1444045',  dbSNP_id=''
WHERE mutation_id = '342'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001257.4:c.193C>T',
HGVS_protein  =  'NP_001248.1:p.R65C',  Location  ='16:83065650',  dbSNP_id=''
WHERE mutation_id = '343'
```

```
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001257.4:c.308C>T',
HGVS_protein    =    'NP_001248.1:p.A103V',    Location    ='16:83065765',
dbSNP_id='rs199539898'  WHERE mutation_id = '344'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001257.4:c.337G>C',
HGVS_protein    =    'NP_001248.1:p.G113R',    Location    ='16:83065794',
dbSNP_id='rs183971768'  WHERE mutation_id = '345'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001257.4:c.736C>T',
HGVS_protein = 'NP_001248.1:p.R246W', Location ='16:83378566', dbSNP_id=''
WHERE mutation_id = '346'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001257.4:c.1099G>C',
HGVS_protein    =    'NP_001248.1:p.E367Q',    Location    ='16:83636197',
dbSNP_id='rs200000145'  WHERE mutation_id = '347'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004490.2:c.1519C>T',
HGVS_protein    =    'NP_004481.2:p.H507Y',    Location    ='2:165349650',
dbSNP_id='rs144301087'  WHERE mutation_id = '348'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_022898.1:c.95A>T',
HGVS_protein = 'NP_075049.1:p.E32V', Location ='14:99724140', dbSNP_id=''
WHERE mutation_id = '351'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '352'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_145243.3:c.205C>T',
HGVS_protein    =    'NP_660286.1:p.H69Y',    Location    ='1:59004762',
dbSNP_id='rs75220198'  WHERE mutation_id = '353'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_145243.3:c.815A>G',
HGVS_protein    =    'NP_660286.1:p.E272G',    Location    ='1:58999918',
dbSNP_id='rs139938730'  WHERE mutation_id = '354'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_145243.3:c.1093G>T',
HGVS_protein    =    'NP_660286.1:p.D365Y',    Location    ='1:58996320',
dbSNP_id='rs77980955'  WHERE mutation_id = '355'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_175733.3:c.460G>A',
HGVS_protein    =    'NP_783860.1:p.V154M',    Location    ='11:7324584',
dbSNP_id='rs78477754'  WHERE mutation_id = '356'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_175733.3:c.712A>G',
HGVS_protein = 'NP_783860.1:p.I238V', Location ='11:7334840', dbSNP_id=''
WHERE mutation_id = '357'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_175733.3:c.1057C>G',
HGVS_protein    =    'NP_783860.1:p.L353V',    Location    ='11:7437285',
dbSNP_id='rs117876446'  WHERE mutation_id = '358'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021076.3:c.2601G>T',
HGVS_protein    =    'NP_066554.2:p.K867N',    Location    ='22:29886230',
dbSNP_id='rs138156220'  WHERE mutation_id = '359'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021076.3:c.2753A>G',
HGVS_protein    =    'NP_066554.2:p.E918G',    Location    ='22:29886382',
dbSNP_id='rs189881592'  WHERE mutation_id = '360'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021076.3:c.1037G>A',
HGVS_protein = 'NP_066554.2:p.R346H', Location ='22:29879517', dbSNP_id=''
WHERE mutation_id = '361'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021076.3:c.119C>T',
HGVS_protein = 'NP_066554.2:p.A40V', Location ='22:29876370', dbSNP_id=''
WHERE mutation_id = '362'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_138966.3:c.1204G>A',
HGVS_protein = 'NP_620416.1:p.G401R', Location ='', dbSNP_id='rs149193005 '
WHERE mutation_id = '363'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_016441.2:c.1217A>G',
HGVS_protein     =     'NP_057525.1:p.N406S',     Location     ='2:36706682',
dbSNP_id='rs139895660'  WHERE mutation_id = '364'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001706.4:c.218A>G',
HGVS_protein = 'NP_001697.2:p.N73S', Location ='3:187449662', dbSNP_id=''
WHERE mutation_id = '365'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021248.2:c.274G>A',
HGVS_protein  =  'NP_067071.1:p.E92K',  Location  ='',  dbSNP_id=''    WHERE
mutation_id = '366'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021248.2:c.401C>T',
HGVS_protein  =  'NP_067071.1:p.T134M',  Location  ='',  dbSNP_id=''    WHERE
mutation_id = '367'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021248.2:c.1598G>A',
HGVS_protein  =  'NP_067071.1:p.R533H',  Location  ='',  dbSNP_id=''    WHERE
mutation_id = '368'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '369'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '370'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_005856.2:c.349G>A',
HGVS_protein = 'NP_005847.1:p.E117K', Location ='7:45222913', dbSNP_id=''
WHERE mutation_id = '371'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_030932.3:c.701T>C',
HGVS_protein = 'NP_112194.2:p.I234T', Location ='13:60554984', dbSNP_id=''
WHERE mutation_id = '372'
GO
UPDATE   dbo.mutation   SET   HGVS_Nucleotide   ='NM_030932.3:c.944   C>T',
HGVS_protein = '', Location ='', dbSNP_id='rs76366906'  WHERE mutation_id =
'373'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '374'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '375'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '376'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '377'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_006940.4:c.1085A>C',
HGVS_protein = 'NP_008871.3:p.Q362P', Location ='12:23757400',
dbSNP_id='rs144670919' WHERE mutation_id = '378'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.67G>T',
HGVS_protein = 'NP_068815.2:p.G23X', Location ='10:13151189', dbSNP_id=''
WHERE mutation_id = '379'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.844A>C',
HGVS_protein = 'NP_068815.2:p.T282P', Location ='10:13164449', dbSNP_id=''
WHERE mutation_id = '380'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.941A>T',
HGVS_protein = 'NP_068815.2:p.Q314L', Location ='10:13166053',
dbSNP_id='rs142812715' WHERE mutation_id = '381'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.552+1delG',
HGVS_protein = '', Location ='10:13154636', dbSNP_id='' WHERE mutation_id
= '382'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.1401+4A>G',
HGVS_protein = '', Location ='10:13169907', dbSNP_id='' WHERE mutation_id
= '383'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.1670A>C',
HGVS_protein = 'NP_068815.2:p.K557T', Location ='10:13178802', dbSNP_id=''
WHERE mutation_id = '384'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_003487.3:c.91G>A',
HGVS_protein = 'NP_003478.1:p.A31T', Location ='17:34147071', dbSNP_id=''
WHERE mutation_id = '385'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein =
'NP_003478.1:p.D386N', Location ='', dbSNP_id='' WHERE mutation_id = '386'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein =
'NP_003478.1:p.R388H', Location ='', dbSNP_id='' WHERE mutation_id = '387'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein =
'NP_003478.1:p.R395Q', Location ='', dbSNP_id='' WHERE mutation_id = '388'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_003487.3:c.1392_1415del24',
HGVS_protein = 'NP_003478.1:p.Arg464_Arg472delinsArg', Location
='17:34171704 - 17:34171727', dbSNP_id='' WHERE mutation_id = '389'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_003487.3:c.1422_1445del',
HGVS_protein = 'NP_003478.1:p.R474_D481del', Location ='17:34171734 -
17:34171757', dbSNP_id='' WHERE mutation_id = '390'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_003487.3:c.717G>A',
HGVS_protein = 'NP_003478.1:p.G239G', Location ='17:34163188',
dbSNP_id='rs144851351' WHERE mutation_id = '391'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '392'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.654T>C',
HGVS_protein = 'NP_068815.2:p.S218S', Location ='10:13160915', dbSNP_id=''
WHERE mutation_id = '393'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.287G>T',
HGVS_protein = 'NP_068815.2:p.R96L', Location ='10:13152394',
dbSNP_id='rs184561087' WHERE mutation_id = '394'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.691_692insAG',
HGVS_protein = 'NP_068815.2:p.Leu231delinsTerTerfs', Location ='10:13160952
- 10:13160953', dbSNP_id='' WHERE mutation_id = '395'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.277G>C',
HGVS_protein = 'NP_068815.2:p.A93P', Location ='10:13152384', dbSNP_id=''
WHERE mutation_id = '396'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.811C>T',
HGVS_protein = 'NP_068815.2:p.R271C', Location ='10:13164416', dbSNP_id=''
WHERE mutation_id = '397'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_025137.3:c.118C>T',
HGVS_protein = 'NP_079413.3:p.Q40X', Location ='15:44955728', dbSNP_id=''
WHERE mutation_id = '398'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '399'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '400'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_025137.3:c.5974C>T',
HGVS_protein = 'NP_079413.3:p.R1992X', Location ='15:44867132',
dbSNP_id='rs200793464' WHERE mutation_id = '401'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_025137.3:c.2198T>G',
HGVS_protein = 'NP_079413.3:p.L733X', Location ='15:44918575', dbSNP_id=''
WHERE mutation_id = '403'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_025137.3:c.2608A>G',
HGVS_protein = 'NP_079413.3:p.I870V', Location ='15:44913969', dbSNP_id=''
WHERE mutation_id = '404'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '405'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '406'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '407'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_025137.3:c.5970C>G',
HGVS_protein = 'NP_079413.3:p.Y1990X', Location ='15:44867136', dbSNP_id=''
WHERE mutation_id = '408'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_025137.3:c.6157G>A',
HGVS_protein = 'NP_079413.3:p.V2053M', Location ='15:44865793',
dbSNP_id='rs149003934' WHERE mutation_id = '409'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '410'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1529A>G',
HGVS_protein = 'NP_004951.1:p.K510R', Location ='16:31202419', dbSNP_id=''
WHERE mutation_id = '411'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1483C>T',
HGVS_protein = 'NP_004951.1:p.R495X', Location ='16:31202373', dbSNP_id=''
WHERE mutation_id = '412'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.760A>G',
HGVS_protein = 'NP_004951.1:p.M254V', Location ='16:31196496', dbSNP_id=''
WHERE mutation_id = '413'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_005866.2:c.304G>C',
HGVS_protein = 'NP_005857.1:p.E102Q', Location ='9:34637265', dbSNP_id=''
WHERE mutation_id = '414'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_013444.3:c.1516C>A',
HGVS_protein = 'NP_038472.2:p.P506T', Location ='X:56591822', dbSNP_id=''
WHERE mutation_id = '415'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_013444.3:c.1490C>A',
HGVS_protein = 'NP_038472.2:p.P497H', Location ='X:56591796', dbSNP_id=''
WHERE mutation_id = '417'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_013444.3:c.1489C>T',
HGVS_protein = 'NP_038472.2:p.P497S', Location ='X:56591795', dbSNP_id=''
WHERE mutation_id = '418'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_013444.3:c.1525C>T',
HGVS_protein = 'NP_038472.2:p.P509S', Location ='X:56591831', dbSNP_id=''
WHERE mutation_id = '419'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_013444.3:c.1573C>T',
HGVS_protein = 'NP_038472.2:p.P525S', Location ='X:56591879', dbSNP_id=''
WHERE mutation_id = '420'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '421'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021980.4:c.476G>T',
HGVS_protein = 'NP_068815.2:p.G159V', Location ='10:13154559', dbSNP_id=''
WHERE mutation_id = '422'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021980.4:c.495C>T',
HGVS_protein  =  'NP_068815.2:Q165X',  Location  ='',  dbSNP_id=''    WHERE
mutation_id = '423'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021980.4:c.1360C>G',
HGVS_protein = 'NP_068815.2:p.Q454E', Location ='10:13169862', dbSNP_id=''
WHERE mutation_id = '425'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '426'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '427'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '428'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021980.4:c.1442C>T',
HGVS_protein = 'NP_068815.2:p.A481V', Location ='10:13174107', dbSNP_id=''
WHERE mutation_id = '429'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_021980.4:c.177G>C',
HGVS_protein  =  'NP_068815.2:p.K59N',  Location  ='10:13152284',  dbSNP_id=''
WHERE mutation_id = '430'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '432'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.35A>C',
HGVS_protein  =  'NP_000445.1:p.D12A',  Location  ='21:33032117',  dbSNP_id=''
WHERE mutation_id = '433'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001917.4:c.113G>A',
HGVS_protein    =    'NP_001908.3:p.R38H',    Location    ='12:109278895',
dbSNP_id='rs3825251'  WHERE mutation_id = '434'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '435'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '436'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '437'
GO
UPDATE  dbo.mutation  SET  HGVS_Nucleotide  ='',  HGVS_protein  =  '',  Location
='', dbSNP_id=''  WHERE mutation_id = '438'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.284T>C',
HGVS_protein    =    'NP_000445.1:p.V95A',    Location    ='21:33039615',
dbSNP_id='rs202198235'  WHERE mutation_id = '439'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.362A>T',
HGVS_protein  =  'NP_000445.1:p.H121L',  Location  ='21:33040788',  dbSNP_id=''
WHERE mutation_id = '440'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.425G>C',
HGVS_protein  =  'NP_000445.1:p.G142A',  Location  ='21:33040851',  dbSNP_id=''
WHERE mutation_id = '441'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.112T>C',
HGVS_protein  =  'NP_001136.1:p.Y38H',  Location  ='14:21161835',  dbSNP_id=''
WHERE mutation_id = '442'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.137A>G',
HGVS_protein  =  'NP_001136.1:p.D46G',  Location  ='14:21161860',  dbSNP_id=''
WHERE mutation_id = '443'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.2:c.1080C>T',
HGVS_protein    =    'NP_004951.1:p.S439S',    Location    ='16:31201374',
dbSNP_id='rs190724342 '  WHERE mutation_id = '444'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.98C>T',
HGVS_protein    =    'NP_003891.1:p.A33V',    Location    ='5:179248034',
dbSNP_id='rs200396166'  WHERE mutation_id = '445'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.457G>A',
HGVS_protein    =    'NP_003891.1:p.V153I',    Location    ='5:179251013',
dbSNP_id='rs145056421'  WHERE mutation_id = '446'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.683C>T',
HGVS_protein    =    'NP_003891.1:p.P228L',    Location    ='5:179252155',
dbSNP_id='rs151191977'  WHERE mutation_id = '447'
```

```
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.702G>A',
HGVS_protein = 'NP_003891.1:p.V234V', Location ='5:179252174', dbSNP_id=''
WHERE mutation_id = '448'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '449'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.783C>T',
HGVS_protein    =    'NP_003891.1:p.H261H',    Location    ='5:179260060',
dbSNP_id='rs145001811'  WHERE mutation_id = '450'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '451'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.961C>T',
HGVS_protein    =    'NP_003891.1:p.R321C',    Location    ='5:179260238',
dbSNP_id='rs140226523'  WHERE mutation_id = '452'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '453'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.1175C>T',
HGVS_protein    =    'NP_003891.1:p.P392L',    Location    ='5:179263445',
dbSNP_id='rs104893941'  WHERE mutation_id = '454'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.1231G>A',
HGVS_protein    =    'NP_003891.1:p.G411S',    Location    ='5:179263501',
dbSNP_id='rs143511494'  WHERE mutation_id = '455'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_003900.4:c.1273G>A',
HGVS_protein = 'NP_003891.1:p.G425R', Location ='5:179263543', dbSNP_id=''
WHERE mutation_id = '456'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '457'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id=''  WHERE mutation_id = '458'
GO
UPDATE            dbo.mutation            SET            HGVS_Nucleotide
='NM_004960.3:c.287291delCCTACinsAT',            HGVS_protein            =
'NP_004951.1:p.Ser96del',  Location  ='16:31488469',  dbSNP_id=''    WHERE
mutation_id = '459'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.667-
678delGGCGGCGGCGGC', HGVS_protein = 'NP_004951.1:p.G223-G226del', Location
='', dbSNP_id=''  WHERE mutation_id = '460'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1449-
1488delCTACCGGGGCCGCGGCGGGGACCGTGGAGGCTTCCGAGGG',      HGVS_protein      =
'NP_004951.1:p.Y485AfsX514', Location ='', dbSNP_id=''  WHERE mutation_id =
'463'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1483delC',
HGVS_protein    =    'NP_004951.1:p.Arg495Glufs',    Location    ='16:31202373',
dbSNP_id=''  WHERE mutation_id = '464'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_004960.3:c.1485delA',
HGVS_protein    =    'NP_004951.1:p.Arg495Arg=fs',    Location    ='16:31202375',
dbSNP_id=''  WHERE mutation_id = '465'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1527insTGGC',
HGVS_protein = 'NP_004951.1:p.K510WfsX517', Location ='', dbSNP_id=''
WHERE mutation_id = '466'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_000454.4:c.412A>G',
HGVS_protein = 'NP_000445.1:p.T138A', Location ='21:33040838', dbSNP_id=''
WHERE mutation_id = '468'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '469'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_000454.4:c.442G>A',
HGVS_protein = 'NP_000445.1:p.G148S', Location ='21:33040868', dbSNP_id=''
WHERE mutation_id = '470'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_001145.4:c.433C>T',
HGVS_protein = 'NP_001136.1:p.R145C', Location ='14:21162156', dbSNP_id=''
WHERE mutation_id = '471'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1484delG',
HGVS_protein = 'NP_004951.1:p.Arg495Glnfs', Location ='16:31202374',
dbSNP_id='' WHERE mutation_id = '472'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '476'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '477'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_001145.4:c.176T>C',
HGVS_protein = 'NP_001136.1:p.L59P', Location ='14:21161899',
dbSNP_id='rs11541242' WHERE mutation_id = '478'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_001145.4:c.250A>G',
HGVS_protein = 'NP_001136.1:p.K84E', Location ='14:21161973',
dbSNP_id='rs17560' WHERE mutation_id = '480'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_002934.2:c.300C>A',
HGVS_protein = ' NP_002925.1:p.His100Gln', Location ='',
dbSNP_id='rs8012891' WHERE mutation_id = '482'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_005022.3:c.211T>G',
HGVS_protein = 'NP_005013.1:p.C71G', Location ='17:4850037', dbSNP_id=''
WHERE mutation_id = '484'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_005022.3:c.341T>C',
HGVS_protein = 'NP_005013.1:p.M114T', Location ='17:4849277', dbSNP_id=''
WHERE mutation_id = '485'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '486'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_005022.3:c.353G>T',
HGVS_protein = 'NP_005013.1:p.G118V', Location ='17:4849265', dbSNP_id=''
WHERE mutation_id = '487'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_000454.4:c.355G>A',
HGVS_protein = 'NP_000445.1:p.V119M', Location ='21:33039686', dbSNP_id=''
WHERE mutation_id = '488'
GO
```

```
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.346C>T',
HGVS_protein = 'NP_000445.1:p.R116C', Location ='21:33039677', dbSNP_id=''
WHERE mutation_id = '489'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_000454.4:c.113G>T',
HGVS_protein = 'NP_000445.1:p.G38V', Location ='21:33036143', dbSNP_id=''
WHERE mutation_id = '490'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.1043G>T',
HGVS_protein = 'NP_031401.1:p.G348V', Location ='1:11082509', dbSNP_id=''
WHERE mutation_id = '492'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.962C>T',
HGVS_protein = 'NP_031401.1:p.A321V', Location ='1:11082428', dbSNP_id=''
WHERE mutation_id = '493'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.-69C>T',
HGVS_protein = '', Location ='1:11072744', dbSNP_id=''  WHERE mutation_id =
'498'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.-66G>T',
HGVS_protein = '', Location ='1:11072747', dbSNP_id=''  WHERE mutation_id =
'500'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.81G>A',
HGVS_protein    =    'NP_031401.1:p.Leu27Leu=',    Location    ='1:11073865',
dbSNP_id=''  WHERE mutation_id = '501'
GO
UPDATE   dbo.mutation   SET   HGVS_Nucleotide   ='NM_007375.3:c.-12-54G>A',
HGVS_protein = '', Location ='1:11073719', dbSNP_id=''  WHERE mutation_id =
'502'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.198T>C',
HGVS_protein    =    'NP_031401.1:p.Ala66Ala=',    Location    ='1:11073982',
dbSNP_id='rs61730366'  WHERE mutation_id = '503'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.312C>T',
HGVS_protein    =    'NP_031401.1:p.Ser104Ser=',    Location    ='1:11076974',
dbSNP_id=''  WHERE mutation_id = '504'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_007375.3:c.411A>G',
HGVS_protein = 'NP_031401.1:p.Lys137', Location ='', dbSNP_id='14:11078798'
WHERE mutation_id = '505'
GO
UPDATE   dbo.mutation   SET   HGVS_Nucleotide   ='NM_007375.3:c.403-80G>A',
HGVS_protein = '', Location ='1:11078710', dbSNP_id=''  WHERE mutation_id =
'506'
GO
UPDATE   dbo.mutation   SET   HGVS_Nucleotide   ='NM_007375.3:c.543+112C>A',
HGVS_protein = '', Location ='1:11079042', dbSNP_id=''  WHERE mutation_id =
'507'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.379G>A',
HGVS_protein = 'NP_001136.1:p.V103I', Location ='14:21162102', dbSNP_id=''
WHERE mutation_id = '508'
GO
UPDATE    dbo.mutation    SET    HGVS_Nucleotide    ='NM_001145.4:c.323A>G',
HGVS_protein = 'NP_001136.1:p.H84R', Location ='14:21162046', dbSNP_id=''
WHERE mutation_id = '509'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1392G>T',
HGVS_protein = 'NP_004951.1:p.M464I', Location ='16:31202162', dbSNP_id=''
WHERE mutation_id = '510'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.52C>T',
HGVS_protein = 'NP_004951.1:p.P18S', Location ='16:31193847',
dbSNP_id='rs144888138' WHERE mutation_id = '511'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.188A>G',
HGVS_protein = 'NP_004951.1:p.N63S', Location ='16:31193983',
dbSNP_id='rs140883211' WHERE mutation_id = '512'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.475 A>T',
HGVS_protein = 'NP_004951.1:p.N159Y', Location ='16:31195669', dbSNP_id=''
WHERE mutation_id = '513'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.491_495 +
1delGAGGTgc', HGVS_protein = 'NP_004951.1:p.G174_G175del', Location ='',
dbSNP_id='' WHERE mutation_id = '515'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide
='NM_004960.3:c.430_447delGGACAGCAGCAAAGCTAT', HGVS_protein =
'NP_004951.1:p.G144_Y149del', Location ='16:31195624 - 16:31195641',
dbSNP_id='' WHERE mutation_id = '517'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '519'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.666_667insGGC',
HGVS_protein = 'NP_004951.1:p.G222_G223insG', Location ='',
dbSNP_id='rs72550890' WHERE mutation_id = '521'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.667_669delGGC',
HGVS_protein = 'NP_004951.1:p.G223del', Location ='', dbSNP_id='rs72550890'
WHERE mutation_id = '523'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.676G>A',
HGVS_protein = 'NP_004951.1:p.G226S', Location ='16:31196412', dbSNP_id=''
WHERE mutation_id = '524'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1147C>T',
HGVS_protein = 'NP_004951.1:p.R383C', Location ='16:31201441', dbSNP_id=''
WHERE mutation_id = '525'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1506dupA',
HGVS_protein = 'NP_004951.1:p.R502fsX15', Location ='', dbSNP_id='' WHERE
mutation_id = '526'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1555C>T',
HGVS_protein = 'NP_004951.1:p.Q519X', Location ='16:31202733', dbSNP_id=''
WHERE mutation_id = '527'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.6C>T', HGVS_protein
= 'NP_004951.1:p.Ala2Ala=', Location ='16:31191541', dbSNP_id='' WHERE
mutation_id = '528'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.147C>A',
HGVS_protein = 'NP_004951.1:p.G49G', Location ='', dbSNP_id='rs741810'
WHERE mutation_id = '529'
GO
```

```
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.153 C>T',
HGVS_protein = 'NP_004951.1:p.G51G', Location ='', dbSNP_id='rs61733962'
WHERE mutation_id = '533'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.222A>G',
HGVS_protein = 'NP_004951.1:p.Gly74Gly=', Location ='16:31195210',
dbSNP_id='' WHERE mutation_id = '534'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.269C>T',
HGVS_protein = 'NP_004951.1:p.Y91Y', Location ='', dbSNP_id='rs73530286'
WHERE mutation_id = '535'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.287C>T',
HGVS_protein = 'NP_004951.1:p.Y97Y', Location ='', dbSNP_id='rs1052352'
WHERE mutation_id = '536'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.504A>T',
HGVS_protein = 'NP_004951.1:p.Gly168Gly=', Location ='16:31195698',
dbSNP_id='' WHERE mutation_id = '537'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.510A>T',
HGVS_protein = 'NP_004951.1:p.Gly170Gly=', Location ='16:31195704',
dbSNP_id='' WHERE mutation_id = '538'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.684C>T',
HGVS_protein = 'NP_004951.1:p.G228G', Location ='16:31196420',
dbSNP_id='rs151073460 ' WHERE mutation_id = '539'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1156C>A',
HGVS_protein = 'NP_004951.1:p.R386R', Location ='', dbSNP_id='rs61733965'
WHERE mutation_id = '540'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1173C>A',
HGVS_protein = 'NP_004951.1:p.P391P', Location ='16:31201600', dbSNP_id=''
WHERE mutation_id = '541'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1464C>T',
HGVS_protein = 'NP_004951.1:p.G488G', Location ='16:31202354',
dbSNP_id='rs150529460' WHERE mutation_id = '542'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_004960.3:c.1566G>A',
HGVS_protein = 'NP_004951.1:p.R522R', Location ='16:31202744',
dbSNP_id='rs138901914 ' WHERE mutation_id = '543'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein =
'NP_001136.1:p.Q(-10)D', Location ='', dbSNP_id='' WHERE mutation_id =
'544'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein =
' NP_001136.1:p.Thr80Ser', Location ='', dbSNP_id='' WHERE mutation_id =
'546'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_001145.4:c.298T>A',
HGVS_protein = ' NP_001136.1:p.Phe100Ile', Location ='',
dbSNP_id='14:21162021' WHERE mutation_id = '548'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein =
' NP_001136.1:p.Asn49Ser', Location ='', dbSNP_id='' WHERE mutation_id =
'550'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '553'
```

```
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '554'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '555'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '556'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '557'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='', HGVS_protein = '', Location
='', dbSNP_id='' WHERE mutation_id = '558'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_001008211.1:c.1107A>G',
HGVS_protein = 'NP_001008212.1:p.L369L', Location ='10:13167526',
dbSNP_id='rs149806984 ' WHERE mutation_id = '559'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_001008211.1:c.1704A>G',
HGVS_protein = 'NP_001008212.1:p.L568L', Location ='10:13178836',
dbSNP_id='' WHERE mutation_id = '560'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.412G>A',
HGVS_protein = 'NP_068815.2:p.T34T', Location ='', dbSNP_id='rs2234968'
WHERE mutation_id = '561'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.433G>A',
HGVS_protein = 'NP_068815.2:p.L41L', Location ='', dbSNP_id='rs11591687'
WHERE mutation_id = '562'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.603T>A',
HGVS_protein = 'NP_068815.2:p.M98K', Location ='', dbSNP_id='rs11258194'
WHERE mutation_id = '563'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.799A>G',
HGVS_protein = 'NP_068815.2:p.E163E', Location ='10:13154572',
dbSNP_id='rs113811959 ' WHERE mutation_id = '564'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.1273C>T',
HGVS_protein = 'NP_068815.2:p.S321S', Location ='10:13166075',
dbSNP_id='rs150381274' WHERE mutation_id = '565'
GO
UPDATE dbo.mutation SET HGVS_Nucleotide ='NM_021980.4:c.1274G>A',
HGVS_protein = 'NP_068815.2:p.E322K', Location ='', dbSNP_id='rs523747'
WHERE mutation_id = '566'
GO

ALTER TABLE dbo.[gene]
ADD [DB_no] int NULL;
GO

UPDATE dbo.gene SET DB_no='1' WHERE gene_id = 'SOD1'
GO
UPDATE dbo.gene SET DB_no='2' WHERE gene_id = 'ALS2'
GO
--UPDATE dbo.gene SET DB_no='3' WHERE gene_id = 'SOD1'
--GO
UPDATE dbo.gene SET DB_no='4' WHERE gene_id = 'SETX'
GO
UPDATE dbo.gene SET DB_no='5' WHERE gene_id = 'SPAST'
```

```
GO
UPDATE dbo.gene SET DB_no='6'  WHERE gene_id = 'FUS'
GO
--UPDATE dbo.gene SET DB_no='7'  WHERE gene_id = 'SOD1'
--GO
UPDATE dbo.gene SET DB_no='8'  WHERE gene_id = 'VAPB'
GO
UPDATE dbo.gene SET DB_no='9'  WHERE gene_id = 'ANG'
GO
UPDATE dbo.gene SET DB_no='10'  WHERE gene_id = 'TARDBP'
GO
UPDATE dbo.gene SET DB_no='11'  WHERE gene_id = 'FIG4'
GO
UPDATE dbo.gene SET DB_no='12'  WHERE gene_id = 'OPTN'
GO
UPDATE dbo.gene SET DB_no='13'  WHERE gene_id = 'ATXN2'
GO
UPDATE dbo.gene SET DB_no='14'  WHERE gene_id = 'VCP'
GO
UPDATE dbo.gene SET DB_no='15'  WHERE gene_id = 'UBQLN2'
GO
UPDATE dbo.gene SET DB_no='16'  WHERE gene_id = 'SIGMAR1'
GO
--UPDATE dbo.gene SET DB_no='17'  WHERE gene_id = 'SOD1'
--GO
UPDATE dbo.gene SET DB_no='18'  WHERE gene_id = 'PFN1'
GO
--UPDATE dbo.gene SET DB_no='19'  WHERE gene_id = 'SOD1'
--GO
UPDATE dbo.gene SET DB_no='20'  WHERE gene_id = 'C9orf72'
GO
UPDATE dbo.gene SET DB_no='21'  WHERE gene_id = 'CHMP2B'
GO
UPDATE dbo.gene SET DB_no='22'  WHERE gene_id = 'UNC13A'
GO
UPDATE dbo.gene SET DB_no='23'  WHERE gene_id = 'DAO'
GO
UPDATE dbo.gene SET DB_no='24'  WHERE gene_id = 'DCTN1'
GO
UPDATE dbo.gene SET DB_no='25'  WHERE gene_id = 'NEFH'
GO
UPDATE dbo.gene SET DB_no='26'  WHERE gene_id = 'PRPH'
GO
UPDATE dbo.gene SET DB_no='27'  WHERE gene_id = 'SQSTM1'
GO
UPDATE dbo.gene SET DB_no='28'  WHERE gene_id = 'TAF15'
GO
UPDATE dbo.gene SET DB_no='29'  WHERE gene_id = 'SPG11'
GO
UPDATE dbo.gene SET DB_no='30'  WHERE gene_id = 'ELP3'
GO
--UPDATE dbo.gene SET DB_no=''  WHERE gene_id = ''
--GO
```

Appendix 38 – Process of generating all possible mutations

ANG_codon.txt

transpose_sequence.pl

repeat_twenty_amino_acids.pl

possibilities.txt

transpose_twenties.pl

possibilities.txt.tr

join_column_notabs.pl

ANG_sequence_twenties.txt.tr + ANG_codon.txt.tr = result3.txt

result3.txt + possiblilities.txt.tr = result4.txt

Example of gene codon AA and NT tranformations using perl script e.g. ANG

ANG codon AA

MVMGLGVLLLVFVLGLGLTPPTLAQDNSRYTHFLTQHYDAKPQGRDDRYCESIMRRRGLT

SPCKDINTFIHGNKRSIKAICENKNGNPHRENLRISKSSFQVTTCKLHGGSPWPPCQYRA

TAGFRNVVVACENGLPVHLDQSIFRRP

ANG codon AA transposed

M	L	T	S	H	D	R
V	V	P	R	Y	D	R
M	F	P	Y	D	R	R
G	V	T	T	A	Y	G
L	L	L	H	K	C	L
G	G	A	F	P	E	T
V	L	Q	L	Q	S	S
L	G	D	T	G	I	P
L	L	N	Q	R	M	C

K	S	P	F	P	F	P
D	I	H	Q	W	R	V
I	K	R	V	P	N	H
N	A	E	T	P	V	L
T	I	N	T	C	V	D
F	C	L	C	Q	V	Q
I	E	R	K	Y	A	S
H	N	I	L	R	C	I
G	K	S	H	A	E	F
N	N	K	G	T	N	R
K	G	S	G	A	G	R
R	N	S	S	G	L	P

ANG codon NT

atggtgatgggcctgggcgttttgttgttggtcttcgtgctgggtctgggtctgacccca

ccgaccctggctcaggataactccaggtacacacacttcctgacccagcactatgatgcc

aaaccacagggccgggatgacagatactgtgaaagcatcatgaggagacggggcctgacc

tcaccctgcaaagacatcaacacatttattcatggcaacaagcgcagcatcaaggccatc

tgtgaaaacaagaatggaaaccctcacagagaaaacctaagaataagcaagtcttctttc

caggtcaccacttgcaagctacatggaggttccccctggcctccatgccagtaccgagcc

acagcggggttcagaaacgttgttgttgcttgtgaaaatggcttacctgtccacttggat

cagtcaattttccgtcgtccgtaa

ANG codon NT transformed

ATG	GGC	GTT	TTG	GTG	CTG	ACC	ACC	CAG	TCC
GTG	CTG	TTG	GTC	CTG	GGT	CCA	CTG	GAT	AGG
ATG	GGC	TTG	TTC	GGT	CTG	CCG	GCT	AAC	TAC

ACA	CAG	AGG	AAC	GCC	GAA	ACC	CCA	GTT	TTG
CAC	GGC	AGA	ACA	ATC	AAC	ACT	TGC	GTT	GAT
TTC	CGG	CGG	TTT	TGT	CTA	TGC	CAG	GTT	CAG
CTG	GAT	GGC	ATT	GAA	AGA	AAG	TAC	GCT	TCA
ACC	GAC	CTG	CAT	AAC	ATA	CTA	CGA	TGT	ATT
CAG	AGA	ACC	GGC	AAG	AGC	CAT	GCC	GAA	TTC
CAC	TAC	TCA	AAC	AAT	AAG	GGA	ACA	AAT	CGT
TAT	TGT	CCC	AAG	GGA	TCT	GGT	GCG	GGC	CGT
GAT	GAA	TGC	CGC	AAC	TCT	TCC	GGG	TTA	CCG
GCC	AGC	AAA	AGC	CCT	TTC	CCC	TTC	CCT	
AAA	ATC	GAC	ATC	CAC	CAG	TGG	AGA	GTC	
CCA	ATG	ATC	AAG	AGA	GTC	CCT	AAC	CAC	

ANG codon

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19
20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23
24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26	26
27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27

28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28
29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29
30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31
32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34	34
35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35
36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36	36
37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37	37
38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38	38
39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39	39
40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41	41
42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42	42
43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43	43
44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44	44
45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45	45
46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46
47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47	47
48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48	48
49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49
50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51	51
52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52	52
53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53	53
54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54	54
55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55	55
56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56	56
57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57
58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58	58
59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59	59
60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60
61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61	61
62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62	62
63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63	63
64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64	64
65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65
66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66	66
67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67	67
68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68	68
69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69	69
70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70	70
71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71	71
72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72	72
73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73	73
74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74
75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75	75
76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76
77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77
78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78
79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79
80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80
81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81

82 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82 82
83 83 83 83 83 83 83 83 83 83 83 83 83 83 83 83 83 83 83 83
84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84 84
85 85 85 85 85 85 85 85 85 85 85 85 85 85 85 85 85 85 85 85
86 86 86 86 86 86 86 86 86 86 86 86 86 86 86 86 86 86 86 86
87 87 87 87 87 87 87 87 87 87 87 87 87 87 87 87 87 87 87 87
88 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88 88
89 89 89 89 89 89 89 89 89 89 89 89 89 89 89 89 89 89 89 89
90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90 90
91 91 91 91 91 91 91 91 91 91 91 91 91 91 91 91 91 91 91 91
92 92 92 92 92 92 92 92 92 92 92 92 92 92 92 92 92 92 92 92
93 93 93 93 93 93 93 93 93 93 93 93 93 93 93 93 93 93 93 93
94 94 94 94 94 94 94 94 94 94 94 94 94 94 94 94 94 94 94 94
95 95 95 95 95 95 95 95 95 95 95 95 95 95 95 95 95 95 95 95
96 96 96 96 96 96 96 96 96 96 96 96 96 96 96 96 96 96 96 96
97 97 97 97 97 97 97 97 97 97 97 97 97 97 97 97 97 97 97 97
98 98 98 98 98 98 98 98 98 98 98 98 98 98 98 98 98 98 98 98
99 99 99 99 99 99 99 99 99 99 99 99 99 99 99 99 99 99 99 99
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10 10
 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9

12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12 12
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13 13
9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14
6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6

14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14 14
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7

ANG codon transformed

1	2	3	5	6	8	9	11	12	14	15
1	2	3	5	6	8	9	11	12	14	15
1	2	4	5	6	8	9	11	12	14	15
1	2	4	5	6	8	9	11	12	14	15
1	2	4	5	7	8	9	11	12	14	15
1	2	4	5	7	8	9	11	12	14	15
1	2	4	5	7	8	10	11	12	14	15
1	2	4	5	7	8	10	11	12	14	15
1	2	4	5	7	8	10	11	13	14	15
1	2	4	5	7	8	10	11	13	14	15
1	2	4	5	7	8	10	11	13	14	16
1	3	4	5	7	8	10	11	13	14	16
1	3	4	5	7	8	10	11	13	14	16
1	3	4	6	7	8	10	11	13	14	16
1	3	4	6	7	8	10	11	13	14	16
1	3	4	6	7	9	10	11	13	14	16
1	3	4	6	7	9	10	11	13	14	16
1	3	4	6	7	9	10	12	13	14	16
1	3	4	6	7	9	10	12	13	14	16
1	3	4	6	7	9	10	12	13	15	16
2	3	4	6	7	9	10	12	13	15	16
2	3	4	6	7	9	10	12	13	15	16
2	3	5	6	7	9	10	12	13	15	16
2	3	5	6	7	9	10	12	13	15	16
2	3	5	6	8	9	10	12	13	15	16
2	3	5	6	8	9	10	12	13	15	16
2	3	5	6	8	9	11	12	13	15	16
2	3	5	6	8	9	11	12	13	15	16
2	3	5	6	8	9	11	12	14	15	16

16	18	19	21	22	24	25	27	28	30	31
17	18	19	21	22	24	25	27	28	30	31
17	18	19	21	22	24	25	27	28	30	31
17	18	20	21	22	24	25	27	28	30	31
17	18	20	21	22	24	25	27	28	30	31
17	18	20	21	23	24	25	27	28	30	31
17	18	20	21	23	24	25	27	28	30	31
17	18	20	21	23	24	26	27	28	30	31
17	18	20	21	23	24	26	27	28	30	31
17	18	20	21	23	24	26	27	29	30	31
17	18	20	21	23	24	26	27	29	30	31
17	18	20	21	23	24	26	27	29	30	32
17	19	20	21	23	24	26	27	29	30	32
17	19	20	21	23	24	26	27	29	30	32
17	19	20	22	23	24	26	27	29	30	32
17	19	20	22	23	24	26	27	29	30	32
17	19	20	22	23	25	26	27	29	30	32
17	19	20	22	23	25	26	27	29	30	32
17	19	20	22	23	25	26	28	29	30	32
17	19	20	22	23	25	26	28	29	30	32
17	19	20	22	23	25	26	28	29	31	32
18	19	20	22	23	25	26	28	29	31	32
18	19	20	22	23	25	26	28	29	31	32
18	19	21	22	23	25	26	28	29	31	32
18	19	21	22	23	25	26	28	29	31	32
18	19	21	22	24	25	26	28	29	31	32
18	19	21	22	24	25	26	28	29	31	32
18	19	21	22	24	25	27	28	29	31	32
18	19	21	22	24	25	27	28	29	31	32

32	34	35	37	38	40	41	43	44	45	47
32	34	35	37	38	40	41	43	44	46	47
33	34	35	37	38	40	41	43	44	46	47
33	34	35	37	38	40	41	43	44	46	47
33	34	36	37	38	40	41	43	44	46	47
33	34	36	37	38	40	41	43	44	46	47
33	34	36	37	39	40	41	43	44	46	47
33	34	36	37	39	40	41	43	44	46	47
33	34	36	37	39	40	42	43	44	46	47
33	34	36	37	39	40	42	43	44	46	47
33	34	36	37	39	40	42	43	45	46	47
33	34	36	37	39	40	42	43	45	46	47
33	34	36	37	39	40	42	43	45	46	48
33	35	36	37	39	40	42	43	45	46	48
33	35	36	37	39	40	42	43	45	46	48
33	35	36	38	39	40	42	43	45	46	48
33	35	36	38	39	40	42	43	45	46	48
33	35	36	38	39	41	42	43	45	46	48
33	35	36	38	39	41	42	43	45	46	48
33	35	36	38	39	41	42	44	45	46	48
33	35	36	38	39	41	42	44	45	46	48
33	35	36	38	39	41	42	44	45	47	48
34	35	36	38	39	41	42	44	45	47	48
34	35	36	38	39	41	42	44	45	47	48
34	35	37	38	39	41	42	44	45	47	48
34	35	37	38	39	41	42	44	45	47	48
34	35	37	38	40	41	42	44	45	47	48
34	35	37	38	40	41	42	44	45	47	48
34	35	37	38	40	41	43	44	45	47	48

48	50	51	53	54	56	57	59	60	61	63
48	50	51	53	54	56	57	59	60	61	63
48	50	51	53	54	56	57	59	60	62	63
49	50	51	53	54	56	57	59	60	62	63
49	50	51	53	54	56	57	59	60	62	63
49	50	52	53	54	56	57	59	60	62	63
49	50	52	53	54	56	57	59	60	62	63
49	50	52	53	55	56	57	59	60	62	63
49	50	52	53	55	56	57	59	60	62	63
49	50	52	53	55	56	58	59	60	62	63
49	50	52	53	55	56	58	59	60	62	63
49	50	52	53	55	56	58	59	61	62	63
49	50	52	53	55	56	58	59	61	62	63
49	50	52	53	55	56	58	59	61	62	64
49	51	52	53	55	56	58	59	61	62	64
49	51	52	53	55	56	58	59	61	62	64
49	51	52	54	55	56	58	59	61	62	64
49	51	52	54	55	56	58	59	61	62	64
49	51	52	54	55	57	58	59	61	62	64
49	51	52	54	55	57	58	59	61	62	64
49	51	52	54	55	57	58	60	61	62	64
49	51	52	54	55	57	58	60	61	62	64
49	51	52	54	55	57	58	60	61	63	64
50	51	52	54	55	57	58	60	61	63	64
50	51	52	54	55	57	58	60	61	63	64
50	51	53	54	55	57	58	60	61	63	64
50	51	53	54	55	57	58	60	61	63	64
50	51	53	54	56	57	58	60	61	63	64
50	51	53	54	56	57	58	60	61	63	64

64	66	67	69	70	72	73	74	76	77	79
64	66	67	69	70	72	73	75	76	77	79
64	66	67	69	70	72	73	75	76	77	79
64	66	67	69	70	72	73	75	76	78	79
65	66	67	69	70	72	73	75	76	78	79
65	66	67	69	70	72	73	75	76	78	79
65	66	68	69	70	72	73	75	76	78	79
65	66	68	69	70	72	73	75	76	78	79
65	66	68	69	71	72	73	75	76	78	79
65	66	68	69	71	72	73	75	76	78	79
65	66	68	69	71	72	74	75	76	78	79
65	66	68	69	71	72	74	75	76	78	79
65	66	68	69	71	72	74	75	77	78	79
65	66	68	69	71	72	74	75	77	78	79
65	66	68	69	71	72	74	75	77	78	80
65	67	68	69	71	72	74	75	77	78	80
65	67	68	69	71	72	74	75	77	78	80
65	67	68	70	71	72	74	75	77	78	80
65	67	68	70	71	72	74	75	77	78	80
65	67	68	70	71	73	74	75	77	78	80
65	67	68	70	71	73	74	75	77	78	80
65	67	68	70	71	73	74	76	77	78	80
65	67	68	70	71	73	74	76	77	78	80
65	67	68	70	71	73	74	76	77	79	80
66	67	68	70	71	73	74	76	77	79	80
66	67	68	70	71	73	74	76	77	79	80
66	67	69	70	71	73	74	76	77	79	80
66	67	69	70	71	73	74	76	77	79	80
66	67	69	70	72	73	74	76	77	79	80

80	82	83	85	86	88	89	90	92	93	95
80	82	83	85	86	88	89	90	92	93	95
80	82	83	85	86	88	89	91	92	93	95
80	82	83	85	86	88	89	91	92	93	95
80	82	83	85	86	88	89	91	92	94	95
81	82	83	85	86	88	89	91	92	94	95
81	82	83	85	86	88	89	91	92	94	95
81	82	84	85	86	88	89	91	92	94	95
81	82	84	85	86	88	89	91	92	94	95
81	82	84	85	87	88	89	91	92	94	95
81	82	84	85	87	88	89	91	92	94	95
81	82	84	85	87	88	90	91	92	94	95
81	82	84	85	87	88	90	91	92	94	95
81	82	84	85	87	88	90	91	93	94	95
81	82	84	85	87	88	90	91	93	94	95
81	82	84	85	87	88	90	91	93	94	96
81	83	84	85	87	88	90	91	93	94	96
81	83	84	85	87	88	90	91	93	94	96
81	83	84	86	87	88	90	91	93	94	96
81	83	84	86	87	88	90	91	93	94	96
81	83	84	86	87	89	90	91	93	94	96
81	83	84	86	87	89	90	91	93	94	96
81	83	84	86	87	89	90	92	93	94	96
81	83	84	86	87	89	90	92	93	94	96
81	83	84	86	87	89	90	92	93	95	96
82	83	84	86	87	89	90	92	93	95	96
82	83	84	86	87	89	90	92	93	95	96
82	83	85	86	87	89	90	92	93	95	96
82	83	85	86	87	89	90	92	93	95	96

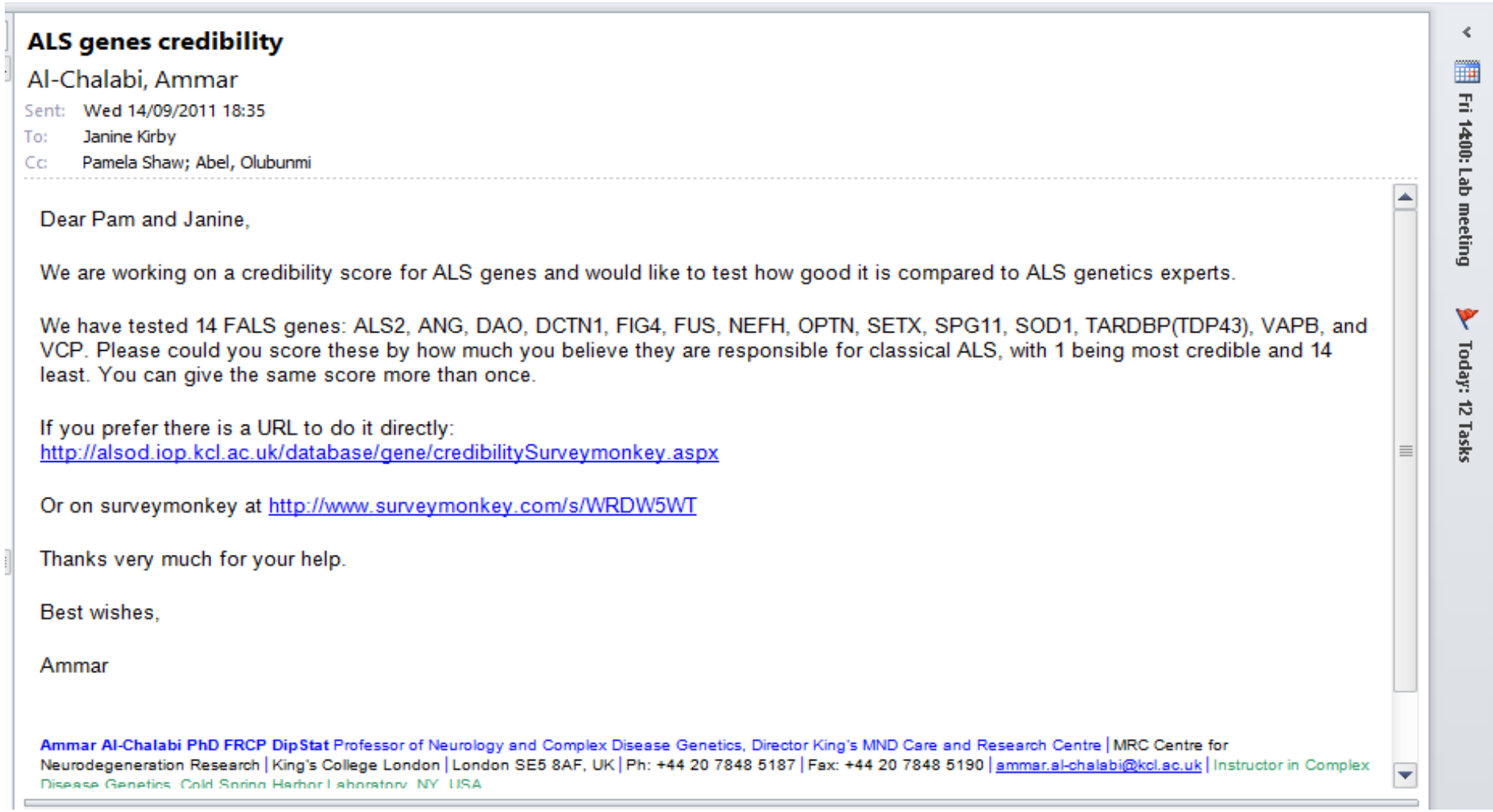
96	98	99	101	102	103	105	106	108	109	111
96	98	99	101	102	104	105	106	108	109	111
96	98	99	101	102	104	105	106	108	109	111
96	98	99	101	102	104	105	107	108	109	111
96	98	99	101	102	104	105	107	108	109	111
96	98	99	101	102	104	105	107	108	110	111
97	98	99	101	102	104	105	107	108	110	111
97	98	99	101	102	104	105	107	108	110	111
97	98	100	101	102	104	105	107	108	110	111
97	98	100	101	102	104	105	107	108	110	111
97	98	100	101	103	104	105	107	108	110	111
97	98	100	101	103	104	105	107	108	110	111
97	98	100	101	103	104	106	107	108	110	111
97	98	100	101	103	104	106	107	108	110	111
97	98	100	101	103	104	106	107	109	110	111
97	98	100	101	103	104	106	107	109	110	111
97	98	100	101	103	104	106	107	109	110	112
97	99	100	101	103	104	106	107	109	110	112
97	99	100	101	103	104	106	107	109	110	112
97	99	100	102	103	104	106	107	109	110	112
97	99	100	102	103	104	106	107	109	110	112
97	99	100	102	103	105	106	107	109	110	112
97	99	100	102	103	105	106	107	109	110	112
97	99	100	102	103	105	106	108	109	110	112
97	99	100	102	103	105	106	108	109	110	112
97	99	100	102	103	105	106	108	109	111	112
98	99	100	102	103	105	106	108	109	111	112
98	99	100	102	103	105	106	108	109	111	112
98	99	101	102	103	105	106	108	109	111	112

112	114	115	117	118	119	121	122	124	125	127
112	114	115	117	118	119	121	122	124	125	127
112	114	115	117	118	120	121	122	124	125	127
112	114	115	117	118	120	121	122	124	125	127
112	114	115	117	118	120	121	123	124	125	127
112	114	115	117	118	120	121	123	124	125	127
112	114	115	117	118	120	121	123	124	126	127
113	114	115	117	118	120	121	123	124	126	127
113	114	115	117	118	120	121	123	124	126	127
113	114	116	117	118	120	121	123	124	126	127
113	114	116	117	118	120	121	123	124	126	127
113	114	116	117	119	120	121	123	124	126	127
113	114	116	117	119	120	121	123	124	126	127
113	114	116	117	119	120	122	123	124	126	127
113	114	116	117	119	120	122	123	124	126	127
113	114	116	117	119	120	122	123	125	126	127
113	114	116	117	119	120	122	123	125	126	127
113	114	116	117	119	120	122	123	125	126	128
113	115	116	117	119	120	122	123	125	126	128
113	115	116	117	119	120	122	123	125	126	128
113	115	116	118	119	120	122	123	125	126	128
113	115	116	118	119	120	122	123	125	126	128
113	115	116	118	119	121	122	123	125	126	128
113	115	116	118	119	121	122	123	125	126	128
113	115	116	118	119	121	122	124	125	126	128
113	115	116	118	119	121	122	124	125	126	128
113	115	116	118	119	121	122	124	125	127	128
114	115	116	118	119	121	122	124	125	127	128
114	115	116	118	119	121	122	124	125	127	128

128	129	130	131	133	134	135	136	137	138	139
128	129	130	131	133	134	135	136	137	138	139
128	129	130	132	133	134	135	136	137	138	139
128	129	130	132	133	134	135	136	137	138	139
128	129	131	132	133	134	135	136	137	138	139
128	129	131	132	133	134	135	136	137	138	139
128	130	131	132	133	134	135	136	137	138	139
128	130	131	132	133	134	135	136	137	138	139
129	130	131	132	133	134	135	136	137	138	140
129	130	131	132	133	134	135	136	137	138	140
129	130	131	132	133	134	135	136	137	139	140
129	130	131	132	133	134	135	136	137	139	140
129	130	131	132	133	134	135	136	138	139	140
129	130	131	132	133	134	135	136	138	139	140
129	130	131	132	133	134	135	137	138	139	140
129	130	131	132	133	134	135	137	138	139	140
129	130	131	132	133	134	136	137	138	139	140
129	130	131	132	133	134	136	137	138	139	140
129	130	131	132	133	135	136	137	138	139	140
129	130	131	132	133	135	136	137	138	139	140
129	130	131	132	134	135	136	137	138	139	
129	130	131	132	134	135	136	137	138	139	
140	140	141	141	142	142	142	143	143	143	144
140	141	141	141	142	142	142	143	143	143	144
140	141	141	141	142	142	142	143	143	143	144
140	141	141	141	142	142	142	143	143	143	144
140	141	141	141	142	142	142	143	143	143	144
140	141	141	141	142	142	142	143	143	144	144
140	141	141	141	142	142	143	143	143	144	144

144	144	145	145	145	146	146	146	147	147	147
144	144	145	145	145	146	146	146	147	147	
144	144	145	145	145	146	146	147	147	147	
144	144	145	145	146	146	146	147	147	147	
144	145	145	145	146	146	146	147	147	147	
144	145	145	145	146	146	146	147	147	147	
144	145	145	145	146	146	146	147	147	147	

Appendix 39 – Sample of letter sent to ALS Experts



ALS genes credibility

Al-Chalabi, Ammar

Sent: Wed 14/09/2011 18:35
To: Janine Kirby
Cc: Pamela Shaw; Abel, Olubunmi

Dear Pam and Janine,

We are working on a credibility score for ALS genes and would like to test how good it is compared to ALS genetics experts.

We have tested 14 FALS genes: ALS2, ANG, DAO, DCTN1, FIG4, FUS, NEFH, OPTN, SETX, SPG11, SOD1, TARDBP(TDP43), VAPB, and VCP. Please could you score these by how much you believe they are responsible for classical ALS, with 1 being most credible and 14 least. You can give the same score more than once.

If you prefer there is a URL to do it directly:
http://alsod.iop.kcl.ac.uk/database/gene/credibilitySurveymonkey.aspx

Or on surveymonkey at http://www.surveymonkey.com/s/WRDW5WT

Thanks very much for your help.

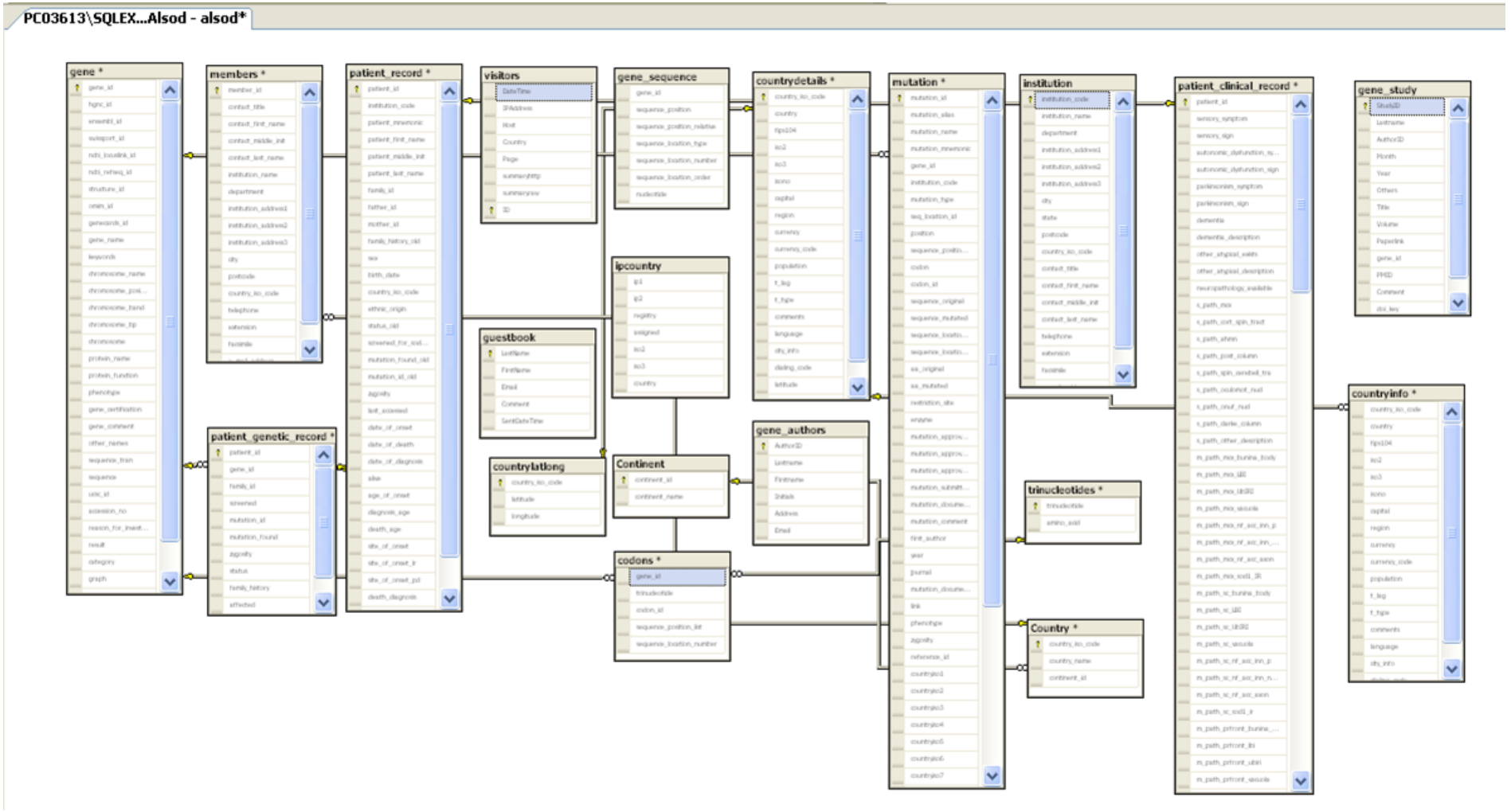
Best wishes,

Ammar

Ammar Al-Chalabi PhD FRCP DipStat Professor of Neurology and Complex Disease Genetics, Director King's MND Care and Research Centre | MRC Centre for Neurodegeneration Research | King's College London | London SE5 8AF, UK | Ph: +44 20 7848 5187 | Fax: +44 20 7848 5190 | ammar.al-chalabi@kcl.ac.uk | Instructor in Complex Disease Genetics, Cold Spring Harbor Laboratory, NY, USA

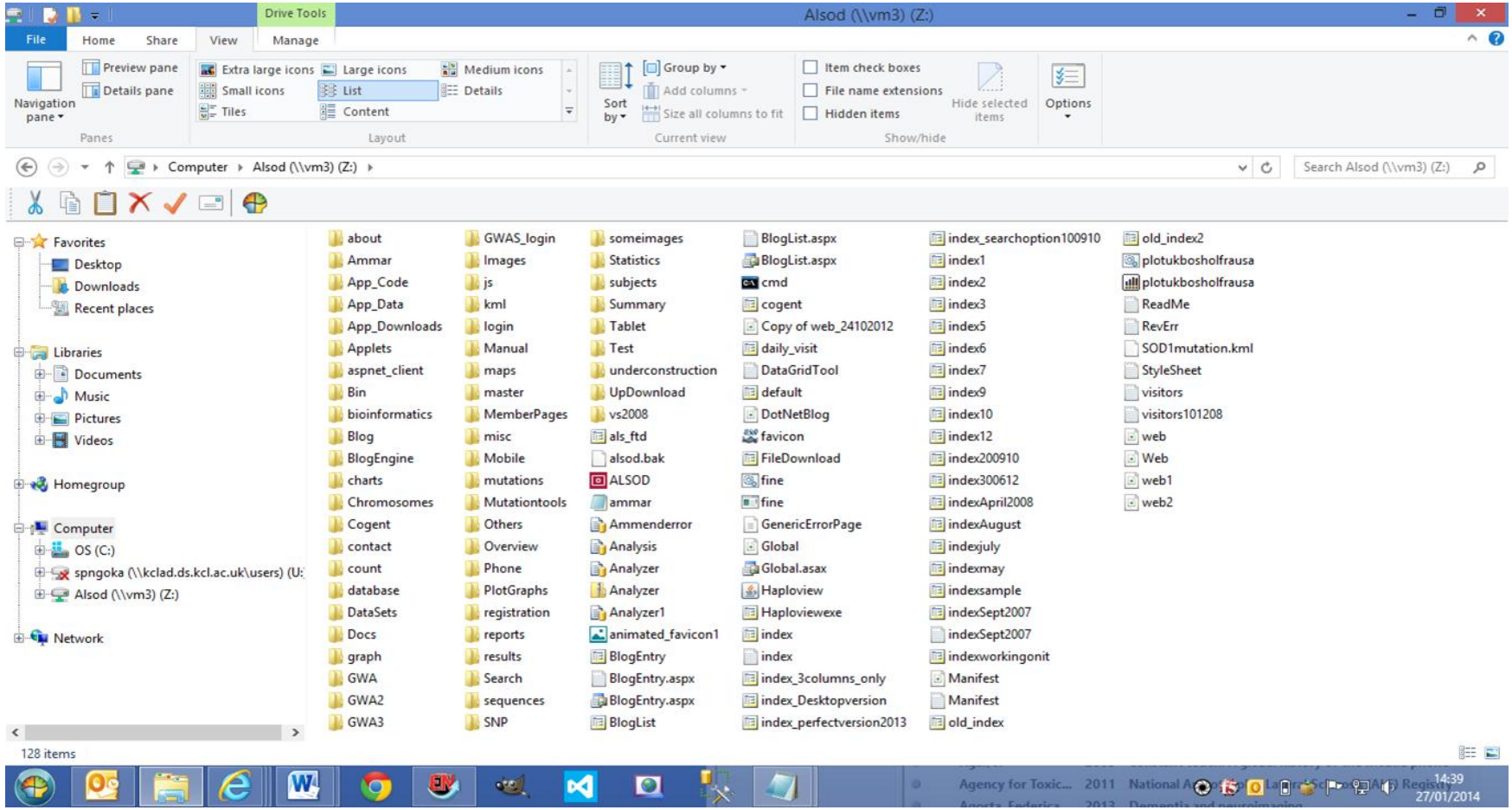
Fri 14:00: Lab meeting
Today: 12 Tasks

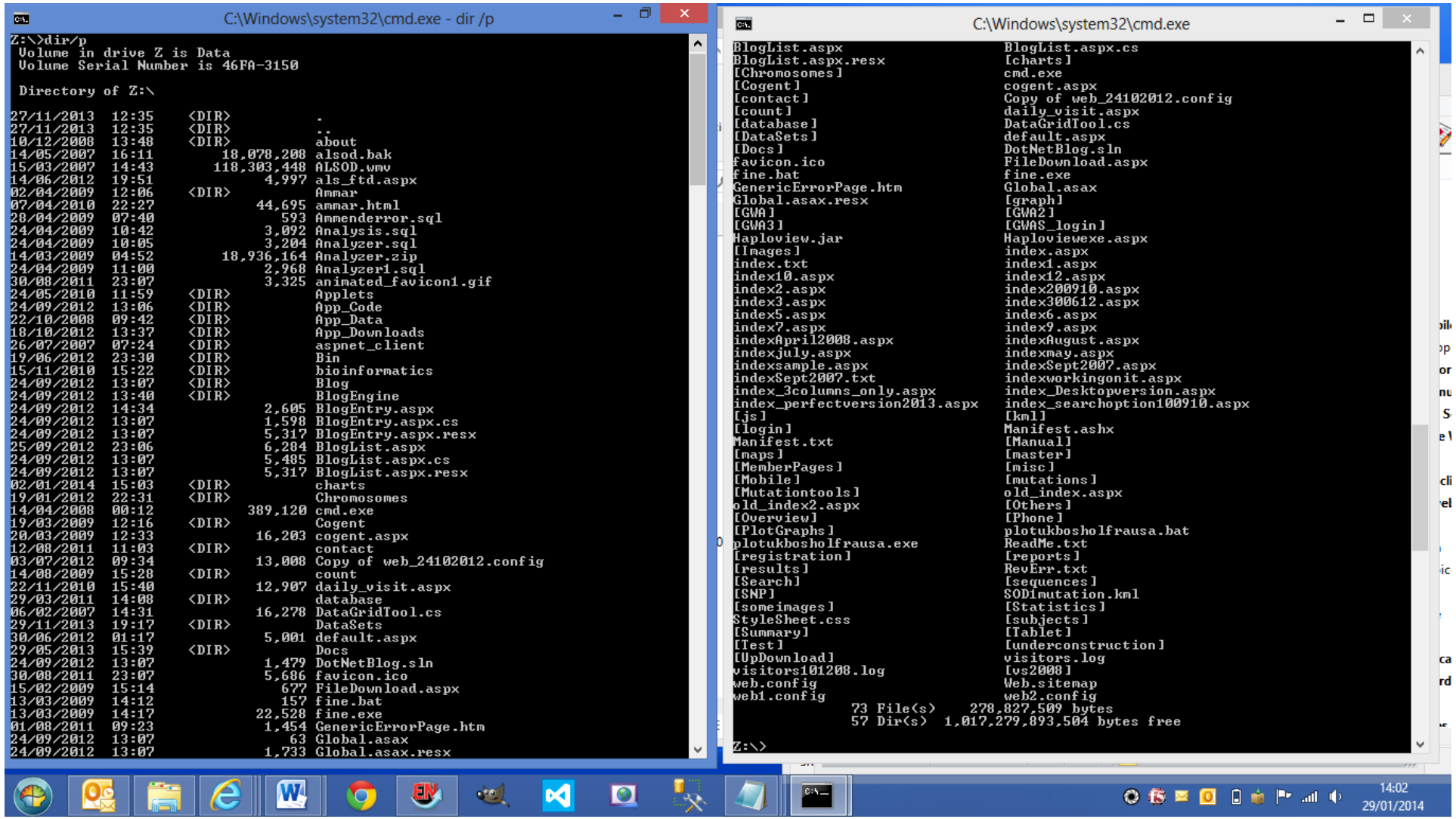
Appendix 40 – Database Schema



Appendix 41 – List of Files

ALSoD Files on KCL Server screen dump taken on 27th January 2014




C:\Windows\system32\cmd.exe - dir /p
Z:\>dir/p
Volume in drive Z is Data
Volume Serial Number is 46FA-3150
Directory of Z:\
27/11/2013 12:35 <DIR> .
27/11/2013 12:35 <DIR> ..
10/12/2008 13:48 <DIR> about
14/05/2007 16:11 18,078,208 alsod.bak
15/03/2007 14:43 118,303,448 ALSOD.wmv
14/06/2012 19:51 4,997 als_ftd.aspx
02/04/2009 12:06 <DIR> Ammar
07/04/2010 22:27 44,695 ammar.html
28/04/2009 07:40 593 Ammenderror.sql
24/04/2009 10:42 3,092 Analysis.sql
24/04/2009 10:05 3,204 Analyzer.sql
14/03/2009 04:52 18,936,164 Analyzer.zip
24/04/2009 11:00 2,968 Analyzer1.sql
30/08/2011 23:07 3,325 animated_favicon1.gif
24/05/2010 11:59 <DIR> Applets
24/09/2012 13:06 <DIR> App_Code
22/10/2008 09:42 <DIR> App_Data
18/10/2012 13:37 <DIR> App_Downloads
26/07/2007 07:24 <DIR> aspnet_client
19/06/2012 23:30 <DIR> Bin
15/11/2010 15:22 <DIR> bioinformatics
24/09/2012 13:07 <DIR> Blog
24/09/2012 13:40 <DIR> BlogEngine
24/09/2012 14:34 2,605 BlogEntry.aspx
24/09/2012 13:07 1,598 BlogEntry.aspx.cs
24/09/2012 13:07 5,317 BlogEntry.aspx.resx
25/09/2012 23:06 6,284 BlogList.aspx
24/09/2012 13:07 5,485 BlogList.aspx.cs
24/09/2012 13:07 5,317 BlogList.aspx.resx
02/01/2014 15:03 <DIR> charts
19/01/2012 22:31 <DIR> Chromosomes
14/04/2008 00:12 389,120 cmd.exe
19/03/2009 12:16 <DIR> Cogent
20/03/2009 12:33 16,203 cogent.aspx
12/08/2011 11:03 <DIR> contact
03/07/2012 09:34 13,008 Copy of web_24102012.config
14/08/2009 15:28 <DIR> count
22/11/2010 15:40 12,907 daily_visit.aspx
29/03/2011 14:08 <DIR> database
06/02/2007 14:31 16,278 DataGridTool.cs
29/11/2013 19:17 <DIR> DataSets
30/06/2012 01:17 5,001 default.aspx
29/05/2013 15:39 <DIR> Docs
24/09/2012 13:07 1,479 DotNetBlog.sln
30/08/2011 23:07 5,686 favicon.ico
15/02/2009 15:14 677 FileDownload.aspx
13/03/2009 14:12 157 fine.bat
13/03/2009 14:17 22,528 fine.exe
01/08/2011 09:23 1,454 GenericErrorPage.htm
24/09/2012 13:07 63 Global.asax
24/09/2012 13:07 1,733 Global.asax.resx
C:\Windows\system32\cmd.exe
BlogList.aspx BlogList.aspx.cs
BlogList.aspx.resx [charts]
[Chromosomes] cmd.exe
[Cogent] cogent.aspx
[contact] Copy of web_24102012.config
[count] daily_visit.aspx
[database] DataGridTool.cs
[DataSets] default.aspx
[Docs] DotNetBlog.sln
favicon.ico FileDownload.aspx
fine.bat fine.exe
GenericErrorPage.htm Global.asax
Global.asax.resx [graph]
[GWA] [GWA2]
[GWA3] [GWAS_login]
Haploview.jar Haplovieweexe.aspx
[Images] index.aspx
index.txt index1.aspx
index10.aspx index12.aspx
index2.aspx index200910.aspx
index3.aspx index300612.aspx
index5.aspx index6.aspx
index7.aspx index9.aspx
indexApril2008.aspx indexAugust.aspx
indexjuly.aspx indexmay.aspx
indexsample.aspx indexSept2007.aspx
indexSept2007.txt indexworkingonit.aspx
index_3columns_only.aspx index_Desktopversion.aspx
index_perfectversion2013.aspx index_searchoption100910.aspx
[js] [kml]
[login] Manifest.ashx
Manifest.txt [Manual]
[maps] [master]
[MemberPages] [misc]
[Mobile] [mutations]
[Mutationtools] old_index.aspx
old_index2.aspx [Others]
[Overview] [Phone]
[PlotGraphs] plotukbosholfrausa.bat
plotukbosholfrausa.exe ReadMe.txt
[registration] [reports]
[results] RevErr.txt
[Search] [sequences]
[SNP] SOD1mutation.kml
[someimages] [Statistics]
StyleSheet.css [subjects]
[Summary] [Tablet]
[Test] [underconstruction]
[UpDownload] visitors.log
visitors101208.log [vs2008]
web.config Web.sitemap
web1.config web2.config
73 File(s) 278,827,509 bytes
57 Dir(s) 1,017,279,893,504 bytes free
Z:\>
14:02
29/01/2014

List of folders

ANG Folder

FUS Folder

OPTN Folder

SETX Folder

SOD1 Folder

TARDBP Folder

DAO Folder

DCTN1 Folder

FIG4 Folder

NEFH Folder

OPTN Folder

SPG11 Folder

UBQLN2 Folder

VAPB Folder

VCP Folder

List of files in each folder e.g ANG folder

ANG_codon.txt

ANG_codon.txt.tr

ANG_possibilities.txt

ANG_possibilities.txt.tr

ANG_sequence.txt

ANG_sequence.txt.tr

ANG_sequence_twenties.txt

ANG_sequence_twenties.txt.tr

ANG_sequence_twenties_with_M.txt

ANG_sequence_twenties_with_M.txt.tr

ANG_sequence_with_M.txt

ANG_sequence_with_M.txt.tr

ANG_SIFT.txt

result1.txt

result2.txt

result3.txt

result4.txt

Perl scripts and data

break_columns.pl

join_columns.pl

join_columns_notabs.pl

repeat_twenty_amino_acids.pl

transpose_sequence.pl

transpose_twenties.pl

Sequence_Worksheet~_ANG.xls

List of Desktop View .aspx files in Web server

/master/MasterPage.master

Index.aspx

/App_Code/AssemblyInfo.cs

/App_Code/ClassDiagram.cd

/App_Code/Global.asax.cs

/App_Code/WebsiteVisitor.cs

/App_Data/ASPNETDB.MDF

/App_Data/aspnetdb_log.LDF

/Applets/FUSspecies.aspx

/Applets/Genespecies.aspx

/Applets/jalviewApplet.jar

/Applets/Jmol-11.0.2.jar

/Applets/OPTNspecies.aspx

/Applets/SOD1.html

/Applets/SOD1Applet.html

/Applets/SOD1species.aspx

/Applets/TARDBP(TDP43)species.aspx

/Applets/Uniref50_sod1.fa

/aspnet_client/system_web/

/Bin/adodb.dll

/Bin/AjaxControlToolkit.dll

/Bin/AjaxControlToolkit.dll.refresh

/Bin/AjaxControlToolkit.pdb

/Bin/Detector.dll

/Bin/Detector.dll.refresh

/Bin/Detector.pdb

/Bin/FiftyOne.Foundation.dll

/Bin/FiftyOne.Foundation.dll.refresh

/Bin/FiftyOne.Foundation.pdb

/Bin/Interop.ChartfxLib.dll

/Bin/Interop.STATCONNECTORCLNTLib.dll

/Bin/Interop.StatConnectorCommonLib.dll

/Bin/Interop.STATCONNECTORSRVLib.dll

/Bin/Interop.StdType.dll

/Bin/KeySortDropDownList.dll

/Bin/KeySortDropDownList.dll.refresh

/Bin/MetaBuilders.WebControls.dll

/Bin/Microsoft.Office.Interop.Access.dll

/Bin/Microsoft.Office.Interop.Excel.dll

/Bin/Microsoft.Office.Interop.FrontPage.dll

/Bin/Microsoft.Office.Interop.Graph.dll

/Bin/Microsoft.Office.Interop.Outlook.dll

/Bin/Microsoft.Office.Interop.OutlookViewCt1.dll

/Bin/Microsoft.Office.Interop.Owc.dll

/Bin/Microsoft.Office.Interop.PowerPoint.dll

/Bin/Microsoft.Office.Interop.Publisher.dll

/Bin/Microsoft.Office.Interop.SmartTag.dll

/Bin/Microsoft.Office.Interop.Visio.dll

/Bin/Microsoft.Office.Interop.Word.dll

/Bin/Microsoft.SqlServer.Smo.dll

/Bin/Microsoft.Vbe.Interop.dll

/Bin/mscomctl.dll

/Bin/msdatasrc.dll

/Bin/office.dll

/Bin/Recaptcha.dll

/Bin/Recaptcha.pdb

/Bin/rscproxy.dll

/Bin/sciproxy.dll

/Bin/StatConnectorClnt.dll

/Bin/stdole.dll

/Bioinformatics/subpsecScore.aspx

/Blog/_UpgradeReport_Files/

/Blog/bin/

/Blog/App_Code/App_Code/AssemblyInfo.cs

/Blog/App_Code/App_Code/ClassDiagram.cd

/Blog/App_Code/App_Code/Global.asax.cs

/Blog/images/separator.jpg

/Blog/BlogEntry.aspx

/Blog/BlogEntry.aspx.cs

/Blog/BlogEntry.aspx.resx

/Blog/BlogList.aspx

/Blog/BlogList.aspx.cs

/Blog/BlogList.aspx.resx

/Blog/comments.xml

/BlogDotNetBlog.sln

/Blog/Global.asax

/Blog/Global.asax.resx

/Blog/RevErr.txt

/Blog/Web.config

/charts/TempImages

/charts/age.aspx

/charts/alsod.xml

/charts/alsod.xml.aspx

/charts/anothermapsod1.aspx

/charts/chart1.aspx

/charts/chartsod1stack.aspx
/charts/codon.aspx
/charts/codon.xls
/charts/data.aspx
/charts/gender.aspx
/charts/googlemap2.aspx
/charts/gwa.aspx
/charts/gwahaploview.aspx
/charts/hitcountstatistics.aspx
/charts/index.aspx
/charts/indexjavascript.aspx
/charts/map.xml
/charts/mutation.aspx
/charts/patients.aspx
/charts/sequence.aspx
/charts/test.aspx
/charts/visitors.html
/charts/visitors_map.html
/charts/visits_comparison.xls
/Chromosomes/chr-1.jpg
/Chromosomes/chr-2.jpg
/Chromosomes/chr-3.jpg
/Chromosomes/chr-4.jpg
/Chromosomes/chr-5.jpg
/Chromosomes/chr-6.jpg
/Chromosomes/chr-7.jpg
/Chromosomes/chr-8.jpg
/Chromosomes/chr-9.jpg
/Chromosomes/chr-10.jpg
/Chromosomes/chr-11.jpg
/Chromosomes/chr-12.jpg
/Chromosomes/chr-13.jpg
/Chromosomes/chr-14.jpg
/Chromosomes/chr-15.jpg
/Chromosomes/chr-16.jpg
/Chromosomes/chr-17.jpg
/Chromosomes/chr-18.jpg
/Chromosomes/chr-19.jpg
/Chromosomes/chr-20.jpg
/Chromosomes/chr-21.jpg
/Chromosomes/chr-22.jpg
/Chromosomes/chr-X.jpg
/Chromosomes/chr-Y.jpg
/Chromosomes/chromo1.aspx
/Chromosomes/chromo2.aspx
/Chromosomes/chromo3.aspx
/Chromosomes/chromo4.aspx
/Chromosomes/chromo5.aspx
/Chromosomes/chromo6.aspx
/Chromosomes/chromo7.aspx
/Chromosomes/chromo8.aspx
/Chromosomes/chromo9.aspx
/Chromosomes/chromo10.aspx
/Chromosomes/chromo11.aspx
/Chromosomes/chromo12.aspx
/Chromosomes/chromo13.aspx
/Chromosomes/chromo14.aspx
/Chromosomes/chromo15.aspx
/Chromosomes/chromo16.aspx
/Chromosomes/chromo17.aspx
/Chromosomes/chromo18.aspx
/Chromosomes/chromo19.aspx
/Chromosomes/chromo20.aspx
/Chromosomes/chromo21.aspx
/Chromosomes/chromo22.aspx
/Chromosomes/chromoX.aspx
/Chromosomes/chromoY.aspx
/Chromosomes/chromoALL.aspx

/Chromosomes/chromo1.jpg
/Chromosomes/chromo2.jpg
/Chromosomes/chromo3.jpg
/Chromosomes/chromo4.jpg
/Chromosomes/chromo5.jpg
/Chromosomes/chromo6jpg
/Chromosomes/chromo7.jpg
/Chromosomes/chromo8.jpg
/Chromosomes/chromo9.jpg
/Chromosomes/chromo10.jpg
/Chromosomes/chromo11.jpg
/Chromosomes/chromo12.jpg
/Chromosomes/chromo13.jpg
/Chromosomes/chromo14.jpg
/Chromosomes/chromo15.jpg
/Chromosomes/chromo16.jpg
/Chromosomes/chromo17.jpg
/Chromosomes/chromo18.jpg
/Chromosomes/chromo19.jpg
/Chromosomes/chromo20.jpg
/Chromosomes/chromo21.jpg
/Chromosomes/chromo22.jpg
/Chromosomes/chromoX.jpg
/Chromosomes/chromoY.jpg
/Chromosomes/chromo1genes.png
/Chromosomes/chromo2genes.png
/Chromosomes/chromo3genes.png
/Chromosomes/chromo4genes.png
/Chromosomes/chromo5genes.png
/Chromosomes/chromo6genes.png
/Chromosomes/chromo7genes.png
/Chromosomes/chromo8genes.png
/Chromosomes/chromo9genes.png
/Chromosomes/chromo10genes.png
/Chromosomes/chromo11genes.png
/Chromosomes/chromo12genes.png
/Chromosomes/chromo13genes.png
/Chromosomes/chromo14genes.png
/Chromosomes/chromo15genes.png
/Chromosomes/chromo16genes.png
/Chromosomes/chromo17genes.png
/Chromosomes/chromo18genes.png
/Chromosomes/chromo19genes.png
/Chromosomes/chromo20genes.png
/Chromosomes/chromo21genes.png
/Chromosomes/chromo22genes.png
/Chromosomes/chromoXgenes.png
/Chromosomes/chromoYgenes.png
/Chromosomes/chromoALLgenes.png
/Chromosomes/chromosome1.ppt
/Chromosomes/chromosome2.ppt
/Chromosomes/chromosome3.ppt
/Chromosomes/chromosome4.ppt
/Chromosomes/chromosome5.ppt
/Chromosomes/chromosome6.ppt
/Chromosomes/chromosome7.ppt
/Chromosomes/chromosome8.ppt
/Chromosomes/chromosome9.ppt
/Chromosomes/chromosome10.ppt
/Chromosomes/chromosome11.ppt
/Chromosomes/chromosome12.ppt
/Chromosomes/chromosome13.ppt
/Chromosomes/chromosome14.ppt
/Chromosomes/chromosome15.ppt
/Chromosomes/chromosome16.ppt
/Chromosomes/chromosome17.ppt
/Chromosomes/chromosome18.ppt
/Chromosomes/chromosome19.ppt

/Chromosomes/chromosome20.ppt
/Chromosomes/chromosome21.ppt
/Chromosomes/chromosome22.ppt
/Chromosomes/chromosomeX.ppt
/Chromosomes/chromosomeY.ppt
/Chromosomes/chromotemplate.ppt
count/counter.xml
/count/Feedback.aspx
/count/hitcounter.aspx
/Chromosomes/Cronbachs.png
/Chromosomes/location_scale.txt
/contact/guestbook.aspx
/contact/contact.aspx
/count/counter.aspx
/count/counter.txt
/count/index.aspx
/count/statistics.aspx
/database/gene/credibilitySurvey.aspx
/database/gene/credibilitySurveymonkey.aspx
/database/gene/submitGeneDetails.aspx
/database/gene/submitGeneDetails2.aspx
/database/mutation/deletionEnterMutationDetails.aspx
/database/mutation/enterIntronMutationDocumentation.aspx
/database/mutation/enterMutationCountry.aspx
/database/mutation/enteMutationDocumentation.aspx
/database/mutation/enterMutationTypeLocation.aspx
/database/mutation/errorPage.aspx
/database/mutation/exonChooseCodon.aspx
/database/mutation/exonChooseNucleotidePosition.aspx
/database/mutation/exonEnterMutationDetails.aspx
/database/mutation/exonMutationDetails.aspx
/database/mutation/intronEnterMutationDetails.aspx
/database/mutation/readyToSubmitMutation.aspx
/database/mutation/verifyIntronMutationDetails.aspx
/database/mutation/verifyMutationDetails.aspx
/database/patient/errorCatch.aspx
/database/patient/Error.aspx
/database/patient/submitPatient000.aspx
/database/patient/submitPatient001.aspx
/database/patient/submitPatient002.aspx
/database/patient/submitPatient003.aspx
/database/patient/submitPatient004.aspx

/database/replication/negativeResults.aspx

/database/replication/submitAuthors.aspx

/database/replication/submitGeneFrequency.aspx

/database/replication/submitReplicatedMutationDetails.aspx

/DataSets/1kg/chr1.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr2.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr3.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr4.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr5.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr6.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr7.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr8.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr9.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr10.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr11.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr12.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr13.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr14.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr15.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr16.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr17.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr18.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr19.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr20.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr21.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chr22.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/1kg/chrX.phase1_release_v3.20101123.snps_indels_svs.minimal2.txt

/DataSets/EVS/chr1.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr2.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr3.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr4.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr5.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr6.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr7.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr8.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr9.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr10.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr11.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr12.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr13.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr14.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr15.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr16.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr17.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr18.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr19.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr20.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr21.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chr22.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chrX.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/chrY.ESP6500SI-V2.snps_indels.txt.minimal2.txt

/DataSets/EVS/ESP6500SI-V2.chr22.snps_indels.txt

/DataSets/5_populations/bos/bos_result_assoc.txt

/DataSets/5_populations/fra/fra_result_assoc.txt

/DataSets/5_populations/hol/hol_result_assoc.txt

/DataSets/5_populations/uk/uk_result_assoc.txt

/DataSets/5_populations/uk/uk_map.txt

/DataSets/5_populations/usa/usa_result_assoc.txt

/DataSets/5_populations/usa/usa_haploview.txt

/DataSets/5_populations/DataPlots/analyse.bat

/DataSets/5_populations/DataPlots/analyse.exe

/DataSets/5_populations/DataPlots/plotall.txt

/DataSets/5_populations/DataPlots/test.txt

/DataSets/5_populations/DataPlots/uk_bos.txt

/DataSets/5_populations/DataPlots/uk_bos_fra.txt

/DataSets/5_populations/DataPlots/uk_bos_fra_usa.txt

/DataSets/5_populations/DataPlots/uk_bos_hol_fra_usa.txt

/DataSets/Animalmodels/AllOrganisms.txt

/DataSets/Animalmodels/MRK_Reference.rpt.txt

/DataSets/Animalmodels/animal_model_upload.xls

/DataSets/Animalmodels/pubmed_references.txt

/GWA/Haploview.aspx

/GWA/gwahaploview.aspx

/GWA/chisquare.aspx

/GWA/gwa.aspx

/GWA/whole.aspx

/GWA/index.aspx

/login/preSign0.aspx

/login/preSign1.aspx

/login/preSign2.aspx

/login/preSign3.aspx

/login/preSign4.aspx

/login/preSign5.aspx

/login/changePassword.aspx

/login/loginAuthenticate.aspx

/login/addUsers.aspx

/login/passwordRecovery.aspx

/login/securePage.aspx

/misc/bioinformatics.aspx

/misc/usefulLinks.aspx

/misc/literature.aspx

/misc/search.aspx

/misc/Sponsors.aspx

/misc/diseseDetails.aspx

/misc/sod1MutationsDiagram.aspx

/misc/Top10.aspx

/misc/contributors.aspx

/misc/labs.aspx

/misc/analysiserror.apx

/misc/Thankyou.aspx

/misc/resources.aspx

/misc/dataDownload.aspx

/misc/FAQs.aspx

/misc/geneFrequency.aspx

/mutation/mutationsFoundCodon.aspx

/mutation/index.aspx

/mutation/mutationsFoundNoGene.aspx

/mutation/mutationsFoundNoType.aspx

/mutation/mutationsFound.aspx

/mutation/mutationsFoundCountry.aspx

/mutation/mutationsFoundMnemonicOnly1.aspx

/mutation/mutationsFoundNoMnemonic.aspx

/mutation/mutationsFoundTypeOnly.aspx

/mutation/mutationsFoundGeneOnly.aspx

/mutation/mutationsFoundMnemonicOnly.aspx

/Mobile/Mutations/mutationsFoundNoType.aspx

/Mobile/Mutations/mutationsFound.aspx

/Mobile/Mutations/Mutation.aspx

/Mobile/index.aspx

/Mobile/MobileMasterPage.master

/Mobile/showPatients.aspx

/Mobile/showMutationDetails.aspx

/Mobile/overview.aspx

/Mobile/publications.aspx

/Mobile/BlogEntry.aspx

/Mobile/BlogList.aspx

/Mobile/feedback.aspx

/Mobile/analysis.aspx

/Mobile/interaction.aspx

/Mobile/comments.txt

/Mobile/Gene.aspx

/Mobile/Mutation.aspx
/Mobile/pathogenicity.aspx
/Mobile/comparison.aspx
/Mobile/credibility.aspx
/Mobile/Animal.aspx
/Mobile/genecattle.aspx
/Mobile/genechicken.aspx
/Mobile/genecimpanzee.aspx
/Mobile/genedog.aspx
/Mobile/generat.aspx
/Mobile/generhesus.aspx
/Mobile/zebrafish.aspx
/Mobile/genemouse.aspx
/Mobile/Animaloverview.aspx
/PlotGraphs/Hol.png
/PlotGraphs/HolUsa.png
/PlotGraphs/HolFraUsa.png
/PlotGraphs/UkBosHolFraUsa.png
/PlotGraphs/BosHolFraUsa.png
/PlotGraphs/Usa.png
/PlotGraphs/FraUsa.png
/PlotGraphs/BosFraUsa.png
/PlotGraphs/UkBosFraUsa.png
/PlotGraphs/Fra.png
/PlotGraphs/BosFra.png
/PlotGraphs/UkBosFra.png
/PlotGraphs/Bos.png
/PlotGraphs/Uk.png
/PlotGraphs/UkBos.png
/PlotGraphs/HolFra.png
/PlotGraphs/BosHol.png
/PlotGraphs/UkHol.png
/PlotGraphs/BosHolUsa.png
/PlotGraphs/BosHolFra.png
/PlotGraphs/UkHolUsa.png
/PlotGraphs/UkHolFra.png
/PlotGraphs/UkBosUsa.png
/PlotGraphs/UkBosHol.png
/PlotGraphs/ukBosHolFra.png
/PlotGraphs/UkBosHolUsa.png
/PlotGraphs/UkHolFraUsa.png
/PlotGraphs/UkFraUsa.png
/PlotGraphs/UkFra.png
/PlotGraphs/BosUsa.png
/PlotGraphs/UkUsa.png
/registration/contributors_detail.aspx
/reports/mutations/showMutations.aspx
/reports/mutations/index.aspx
/reports/allMutations.aspx
/reports/caseSummaries.aspx
/reports/caseSummaries.aspx
/reports/allPatients.aspx
/reports/reportSummary.aspx
/results/showPatients.aspx
/results/showMutationsDetails.aspx
/Search/SearchResultPage.aspx
/Search/searchWeb.aspx
/Statistics/fullindexsample.aspx
/Statistics/mutation.aspx
/Statistics/search.aspx
/Statistics/index2.aspx
/Statistics/study2.aspx
/Statistics/protein.aspx
/Statistics/study.aspx
/Statistics/gene.aspx
/Statistics/index.aspx
/Statistics/interaction.aspx
/Statistics/report.aspx

/Statistics/test.aspx

/Statistics/credibility.aspx

/Statistics/pathogenicity.aspx

/Statistics/statstics.aspx

/Statistics/analysis.aspx

/subjects/index.aspx

/subjects/searchPatientsResults.aspx

/subjects/ searchPatientsResults2.aspx

/subjects/ searchPatientsResults3.aspx

/Summary/summary.aspx

/master/MasterPage.aspx

/Overview/mutation.aspx

/Overview/interaction.aspx

/Overview/gene.aspx

/Overview/genezebrafish.aspx

/Overview/genecattle.aspx

/Overview/genechicken.aspx

/Overview/genechimpanzee.aspx

/Overview/genecattle.aspx

/Overview/generhesus.aspx

/Overview/genedog.aspx

/Overview/generat.aspx

/Overview/genemouse.aspx

/Overview/protein.aspx

/Overview/search.aspx

/Overview/study.aspx

/Overview/index.aspx

/Images/Forestplots/APEX1.png

/Images/Forestplots/PON1.png

/Images/Forestplots/PON2.png

/Images/Forestplots/PON3.png

/Images/Forestplots/ANG.png

/Images/Forestplots/HFE.png

/Images/Forestplots/VEGFA.png

/Images/Forestplots/APOE.png

/Images/Forestplots/UNIC13A.png

/Images/Forestplots/C9orf72.png

/Images/Forestplots/GWA_9p21-2_rs3849942

/Images/blog.ppt

/Images/mobile_banner.ppt

/Images/ALS_FTD.png

/Images/alsalogo.png

/Images/ALSGene.png

/Images/ALSmutationdatabase.png

/Images/alsodlogo.png

/Images/alsodlogo_big.png

/Images/alsodlogo_small.png

/Images/blog.png

/Images/cookie.png

/Images/footer-grey.png

/Images/Googlescholar.png

/Images/GWAS_Fogh.png

/Images/HGMD.png

/Images/Jalview.png

/Images/joint_project_banner.png

/Images/logo_blog.png

/Images/logo_dog.png

/Images/logo_human.png

/Images/mobile_banner.png

/Images/newlogo.png

/Images/Panther.png

/Images/PolyPhen.png

/Images/Pubmed.png

/Images/road-hazard.png

/Images/road-information.png

/Images/road-works.png

/Images/Sift.png

/Images/suggest-post.png
/Images/transparent-bg.png
/Images/Uniprot.png
/Images/my_sb_logo.gif
/Images/minus.gif
/Images/plus.gif
/Images/my_als_us_logo.gif
/Images/my_mnda_uk_logo.gif
/Images/new_banner.gif
/Images/logo.gif
/Images/alscanadalogo.gif
/Images/arrow_blue_v_Variation_5.gif
/Images/bkg-route-no.gif
/Images/collapse.gif
/Images/envelope.gif
/Images/expand.gif
/Images/leftWaveTop.gif
/Images/LsqBltBlue.gif
/Images/LsqBltBlueon.gif
/Images/menu-bullet.gif
/Images/minus.gif
/Images/File.gif
/Images/Folder.gif
/Images/home.gif
/Images/iop.gif
/Images/jpbuttonbg.gif
/Images/alscanada.jpg
/Images/dnaDoubleHelix.jpg
/Images/sod1ProteinPicture.jpg
/Images/brainDiagram.jpg
/Images/dna4.jpg
/Images/ingenuityLogo.jpg
/Images/banner.jpg
/Images/brainDiagram.jpg
/Images/dnaDoubleHelix.jpg
/Images/falsconnectlogo.jpg
/Images/logo_animal.jpg
/Images/logo_cattle.jpg
/Images/logo_chicken.jpg
/Images/logo_chimpanzee.jpg
/Images/logo_comparison.jpg
/Images/logo_credibility.jpg
/Images/logo_gene.jpg
/Images/logo_mutation.jpg
/Images/logo_patho.jpg
/Images/logo_rat.jpg
/Images/logo_rhesus.jpg
/Images/logo_zebrafish.jpg
/Images/logo_blog.png
/Images/logo_dog.png
/Images/logo_human.png
/Images/logo_mouse.png
/Images/cattle.jpg
/Images/chicken.jpg
/Images/chimpanzee.jpg
/Images/rat.jpg
/Images/rhesus.jpg
/Images/zebrafish.jpg
/Images/blog.png
/Images/dog.png
/Images/human.png
/Images/mouse.png
/Images/logo2009.png
/Images/logo2010.png
/Images/separator.jpg
/Images/microarray.jpg
/Images/tcsomeBackground4.jpg
/Images/logo_test-tubes.jpg

/Images/OpenFolder.gif

/Images/rtArwGrey.gif

/Images/rtArwLtBlue.gif

/Images/sqBltGrey.gif

/Images/tab-background.gif

/Images/textSizeSlice.gif

/Images/Up.gif

/Images/worldmap.gif

/js/jquery-1.3.2.min.js

/js/jquery-ui-1.7.2.custom.min.js

/js/tabcontent.js

/login/addUsers.aspx

/registration/contributors_detail.aspx

/subjects/index.aspx

/subjects/searchPatientsResults.aspx

/subjects/searchPatientsResults2.aspx

/subjects/searchPatientsResults3.aspx

/Summary/summary.aspx

/UpDownload/FileDownload.aspx

/UpDownload/index.aspx

Spreadsheets

- animalmodels.xls – To update animal model table
- ALSoD_update_july_2013.xls – To update mutation and patient data
- Survey_ranking_credibility_genes.xls

Poster Presentations

- ALSOD_Advert.xls 19th August 2009
- ALSOD_MNDA_Poster 13th October 2009
- ALSOD_MNDA_Poster(ammended) 22nd March 2010
- ALSOD_ABN_Poster(ammended) 15th April 2010
- ALSOD_Advert2010 11th September 2010
- ALSOD_WELLCOMETRUST_Poster_ammended 6th December 2010
- ALSOD_MNDA_2010_Poster 8th December 2010
- ALSOD_MRC_Poster25082011 30th August 2011
- ALSOD_MNDA_Poster2012 26th November 2012
- ALSOD_Advert2013 4th September 2013

Oral Presentations

- ALSOD PRESENTATION2008. ppt
- Presentation231110.ppt November 2010
- Showcase_Credibility. ppt Novemeber 2011

- Showcase November 2012
- Departmental talk on ALS database January 2013
- ALSoD_MNDA2013. ppt December 2013
- Departmental talk on ALS database February 2014

Abstracts submitted and approved for presentations

- Showcase_mobile_friendly.doc words November 2012
- Showcase_abstract_and_form_2011.doc words November 2011
- Showcase_Abstract proforma.doc
- ALSoD_abstract_MNDA_Italy2013.doc

List of Python script on population frequency

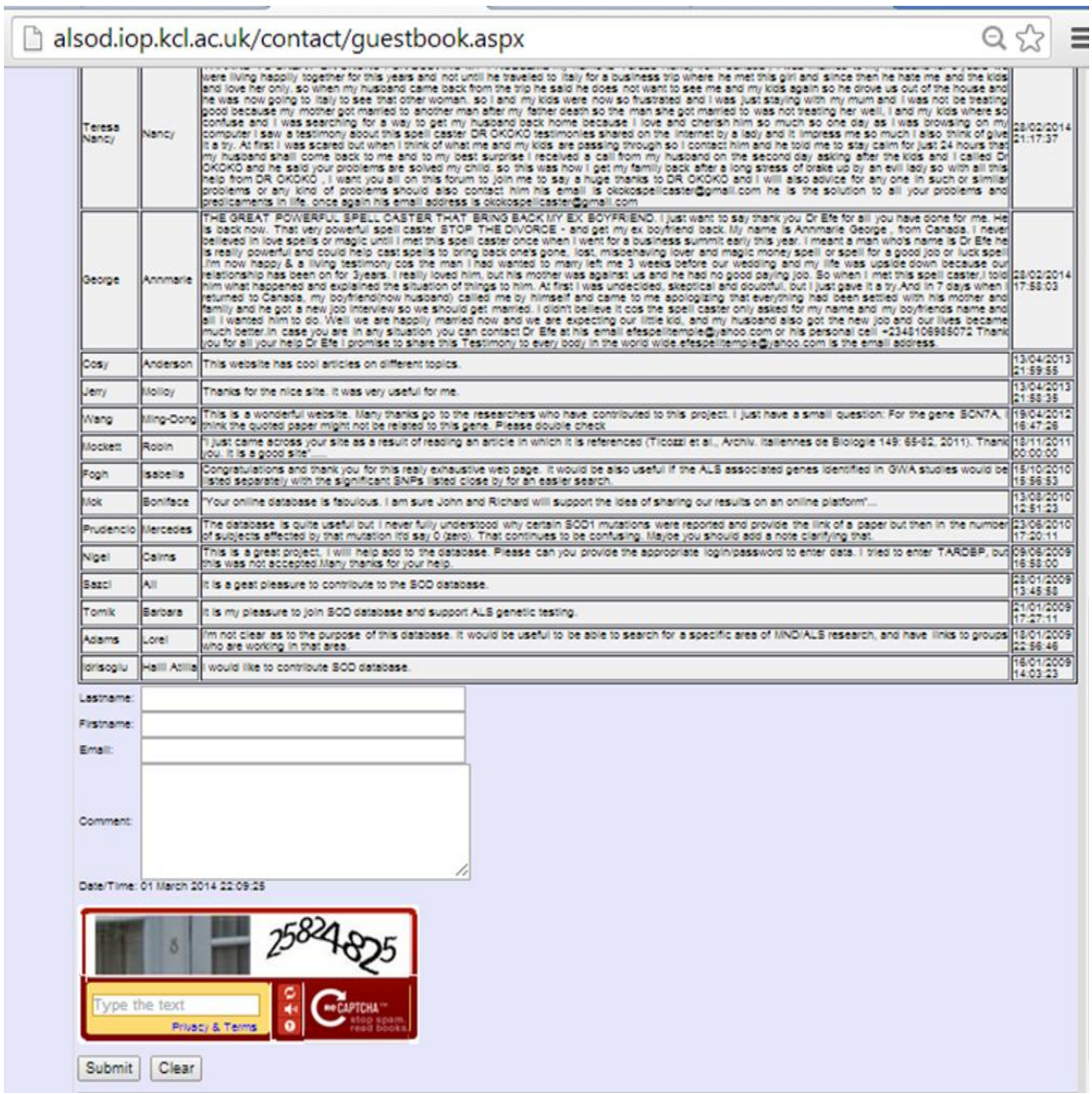
snp_match.py

List of textfiles imported into database

countries, gwas,1000genome,EVS

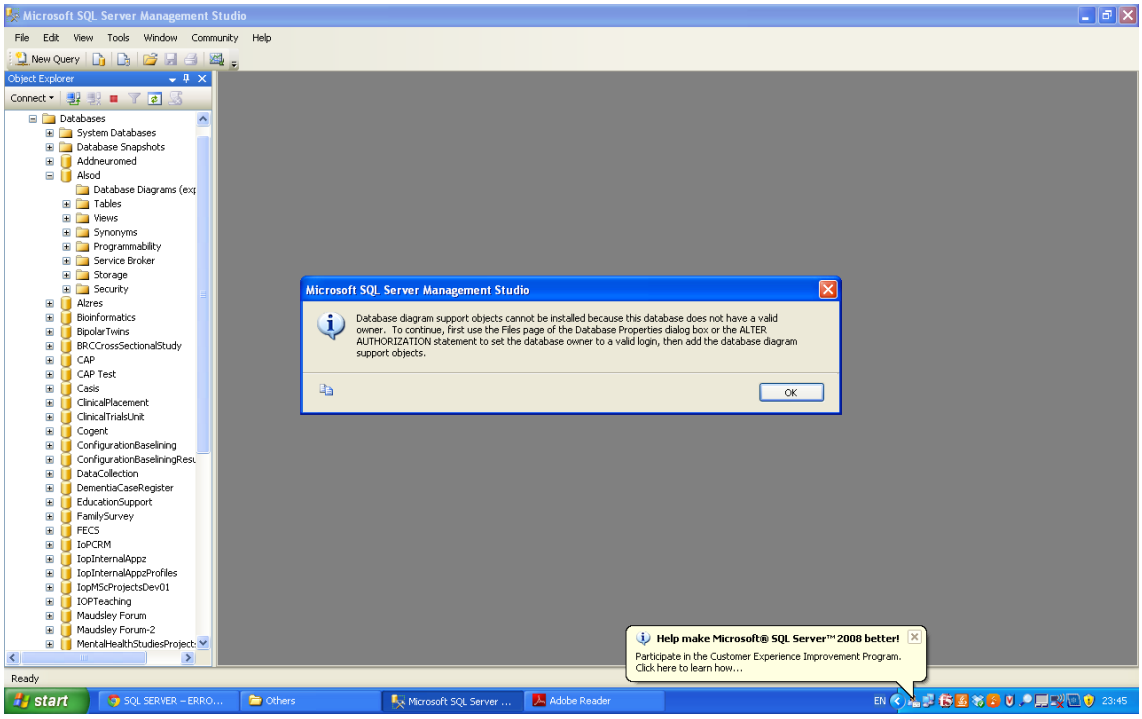
Appendix 42 - Recaptcha


Google
alsonlinedatabase@gmail.com
reCAPTCHA
alsod.iop.kcl.ac.uk
HOME
WHAT IS reCAPTCHA
GET reCAPTCHA
MY ACCOUNT
MY SITES
EMAIL PROTECTION
RESOURCES
Domain Name:
alsod.iop.kcl.ac.uk
This is a global key. It will work across all domains.
Public Key:
6LdkZO8SAAAAAJ4pGz55bXbCihqGhQcZqpvj9Z7X
Use this in the JavaScript code that is served to your users
Private Key:
6LdkZO8SAAAAAIgs_oCXEYZNMPp0x8wbBURtfm-g
Use this when communicating between your server and our server. Be sure to keep it a secret.
Delete these keys
Resources:
reCAPTCHA plugins and libraries
reCAPTCHA API Documentation
Now what?
SUBSCRIBE to important reCAPTCHA service announcements.
Install reCAPTCHA on your site. This is done in two parts. First, you need to add some HTML that displays the reCAPTCHA widget. Second, you need to configure your form to contact our servers to verify reCAPTCHA solutions. Here are specific instructions for: PHP and MediaWiki. For other environments, visit our resources page.
If you need help, post your questions in the reCAPTCHA forum.
Blog | About Us | Contact | FAQ | Terms
© 2014 Google, all rights reserved.


alsod.iop.kcl.ac.uk/contact/guestbook.aspx
Teresa Nancy
Nancy
were living happily together for this years and not until he traveled to Italy for a business trip where he met this girl and since then he hate me and the kids and love her only, so when my husband came back from the trip he said he does not want to see me and my kids again so he drove us out of the house and he was now going to Italy to see that other woman, so i and my kids were now so frustrated and I was just staying with my mum and I was not be treating good because my mother got married to another man after my father death so the man she got married to was not treating her well, I and my kids where so confuse and I was searching for a way to get my husband back home because I love and cherish him so much so one day as I was browsing on my computer I saw a testimony about this spell caster DR OKOKO testimonies shared on the internet by a lady and it impress me so much I also think of give it a try. At first I was scared but when I think of what me and my kids are passing through so I contact him and he told me to stay calm for just 24 hours that my husband shall come back to me and to my best surprise I received a call from my husband on the second day asking after the kids and I called Dr OKOKO and he said your problems are solved my child, so this was how I get my family back after a long stress of broke up by an evil lady so with all this help from DR OKOKO , I want you all on this forum to join me to say a huge thanks to DR OKOKO and I will also advice for any one in such or similar problems or any kind of problems should also contact him his email is okokospellcaster@gmail.com he is the solution to all your problems and predicaments in life, once again his email address is okokospellcaster@gmail.com
28/02/2014 21:17:37
George
Annmarie
THE GREAT POWERFUL SPELL CASTER THAT BRING BACK MY EX BOYFRIEND. I just want to say thank you Dr Efe for all you have done for me. He is back now. That very powerful spell caster STOP THE DIVORCE - and get my ex boyfriend back. My name is Annmarie George , from Canada. I never believed in love spells or magic until I met this spell caster once when I went for a business summit early this year. I meant a man who's name is Dr Efe he is really powerful and could help cast spells to bring back one's gone, lost, misbehaving lover and magic money spell or spell for a good job or luck spell. I'm now happy & a living testimony cos the man I had wanted to marry left me 3 weeks before our wedding and my life was upside down because our relationship has been on for 3years. I really loved him, but his mother was against us and he had no good paying job. So when I met this spell caster,I told him what happened and explained the situation of things to him. At first I was undecided, skeptical and doubtful, but I just gave it a try. And in 7 days when I returned to Canada, my boyfriend(now husband) called me by himself and came to me apologizing that everything had been settled with his mother and family and he got a new job interview so we should get married. I didn't believe it cos the spell caster only asked for my name and my boyfriends name and all I wanted him to do. Well we are happily married now and we are expecting our little kid, and my husband also got the new job and our lives became much better.In case you are in any situation you can contact Dr Efe at his email efespelltemple@yahoo.com or his personal cell +2348106985072 Thank you for all your help Dr Efe I promise to share this Testimony to every body in the world wide efespelltemple@yahoo.com is the email address.
28/02/2014 17:58:03
Cosy
Anderson
This website has cool articles on different topics.
13/04/2013 21:59:56
Jerry
Molloy
Thanks for the nice site. It was very useful for me.
13/04/2013 21:55:35
Wang
Ming-Dong
This is a wonderful website. Many thanks go to the researchers who have contributed to this project. I just have a small question: For the gene BCN7A, I think the quoted paper might not be related to this gene. Please double check
19/04/2012 16:47:26
Mockett
Robin
"I just came across your site as a result of reading an article in which it is referenced (Tricozzi et al., Archiv. Italiennes de Biologie 149: 65-82, 2011). Thank you. It is a good site"....
18/11/2011 00:00:00
Fogh
Isabella
Congratulations and thank you for this realy exhaustive web page. It would be also useful if the ALS associated genes identified in GWA studies would be listed separately with the significant SNPs listed close by for an easier search.
15/10/2010 15:56:53
Mok
Boniface
"Your online database is fabulous. I am sure John and Richard will support the idea of sharing our results on an online platform"....
13/08/2010 12:51:23
Prudencio
Mercedes
The database is quite useful but I never fully understood why certain SOD1 mutations were reported and provide the link of a paper but then in the number of subjects affected by that mutation it's say 0 (zero). That continues to be confusing. Maybe you should add a note clarifying that.
23/06/2010 17:20:11
Nigel
Cairns
This is a great project. I will help add to the database. Please can you provide the appropriate login/password to enter data. I tried to enter TARDBP, but this was not accepted.Many thanks for your help.
09/06/2009 16:58:00
Bazci
Ali
It is a geat pleasure to contribute to the SOD database.
23/01/2009 13:45:58
Tomik
Barbara
It is my pleasure to join SOD database and support ALS genetic testing.
21/01/2009 17:27:11
Adams
Lorel
I'm not clear as to the purpose of this database. It would be useful to be able to search for a specific area of MND/ALS research, and have links to groups who are working in that area.
19/01/2009 22:56:46
Idrisoglu
Halil Atilla
I would like to contribute SOD database.
16/01/2009 14:03:23
Lastname:
Firstname:
Email:
Comment:
Date/Time: 01 March 2014 22:09:26
25824825
Type the text
Privacy & Terms
reCAPTCHA stop spam. read books.
Submit
Clear

Appendix 43 – Errors

Error 1

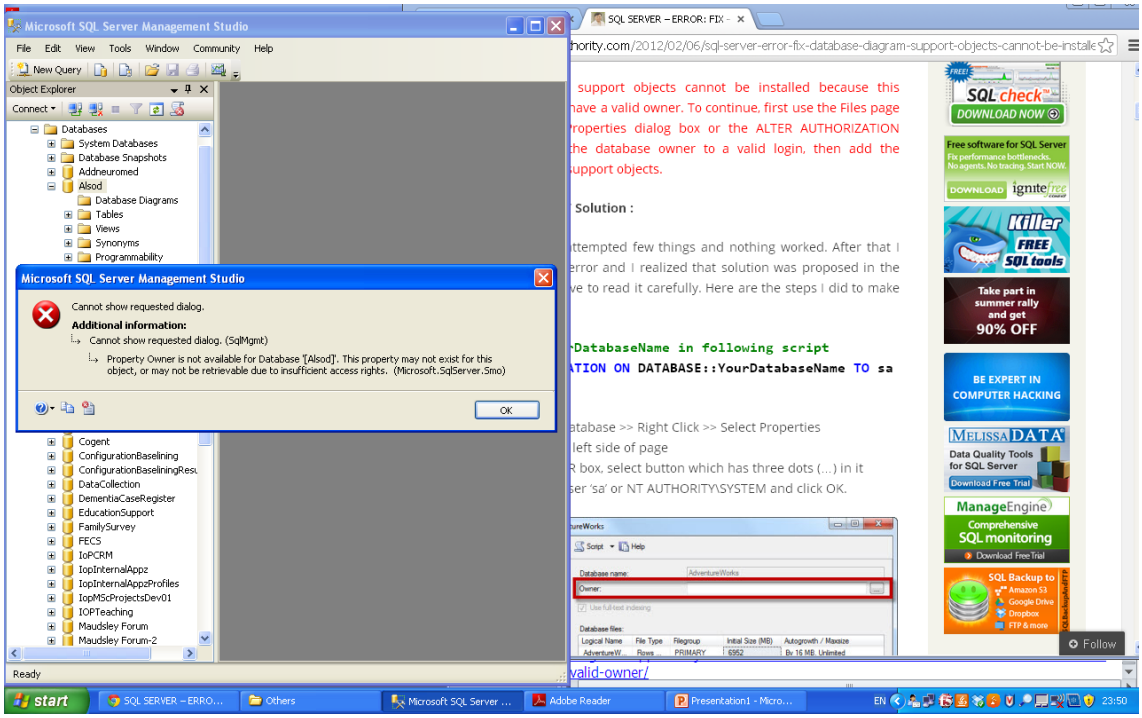


Solution 1:

Database diagram support objects cannot be installed because this database does not have a valid owner. To continue, first use the Files page of the Database Properties dialog box or the ALTER AUTHORIZATION statement to set the database owner to a valid login, then add the database diagram support objects.

Solution from blog: http://blog.sqlauthority.com/2012/02/06/sql-server-error-fix-database-diagram-support-objects-cannot-be-installed-because-this-database-does-not-have-a-valid-owner/

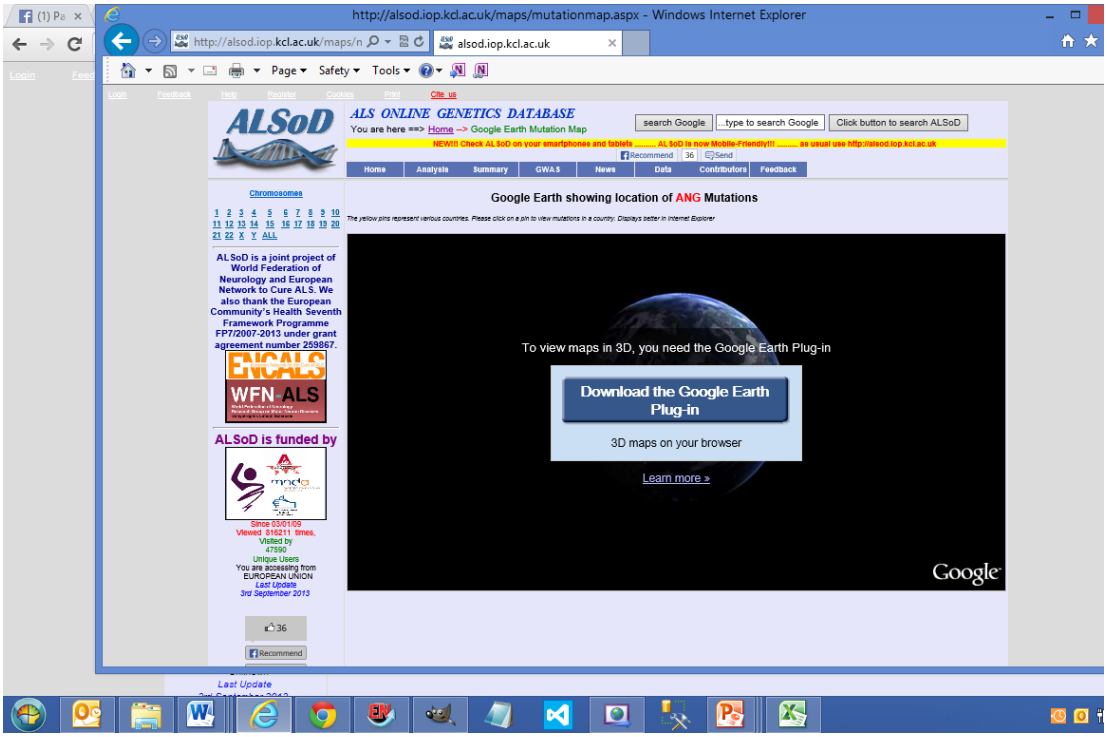
Error 2



Solution 2:

Same as solution 1 where the owner is changed appropriately.

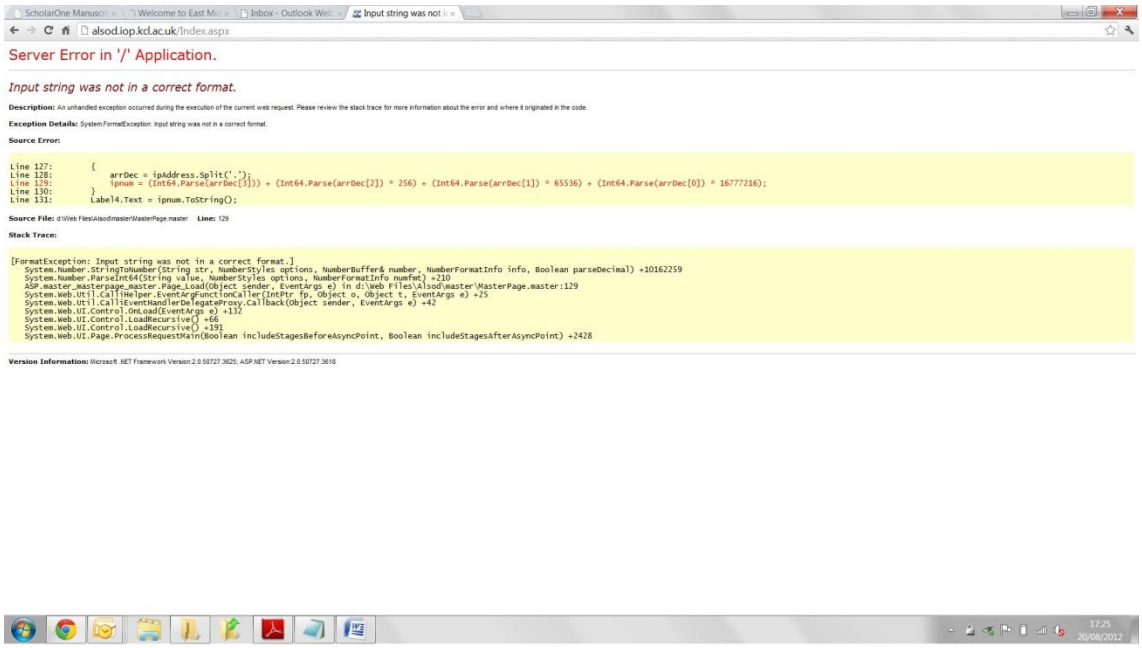
Error 3



Solution 3:

If google map does not appear immediately but requests the user to download the plug-in, then the user can click on the download button. This is because the user has not got the java plugin required to view the map on his/her machine.

Error 4



Solution 4:

```
<customErrors mode="RemoteOnly"
defaultRedirect="GenericErrorPage.htm">
    <error  statusCode="403" redirect="NoAccess.htm"/>
    <error statusCode="404" edirect="FileNotFound.htm"/>
  </customErrors>
```

Where

```
 <customErrors mode="Off"
defaultRedirect="GenericErrorPage.htm">
```

any error does not show the error encountered but displays the default error page.

Error 5

```
    1.  Msg 4861, Level 16, State 1, Line 2
Cannot bulk load because the file "C:\Documents and Settings\spngoka\Desktop\Alsod
stuff\DataSets\uk\uk_result_assoc.txt" could not be opened. Operating system error code
3(error not found).
```

Solution 5:

Check to ensure 'C:\Documents and Settings\spngoka\Desktop\Alsod stuff\DataSets\uk\uk_result_assoc.txt ' exists at the specified location on the SQL Server. Note that the path is accessed from the server rather than the client. If the file resides on the client machine, you'll

need to create a share and assign file/share permissions to the SQL Server service in order to BULK INSERT from a network location. The best option was to copy the file to the webserver \\vmw\Alsod\ so it could be accessed directly while bulk inserting.

Solution blog website:

http://www.eggheadcafe.com/forumarchives/SQLServerprogramming/Jan2006/post26056833.asp

I finally, copied C:\Documents and Settings\spngoka\Desktop\Alsod stuff\DataSets\uk\uk_result_assoc.txt into r: drive and acessed it as \\vm3\DataSets\uk\uk_result_assoc.txt

Error 6

```
2. The statement has been terminated.

Msg 9002, Level 17, State 2, Line 1
```

```
The transaction log for database 'Alsod' is full.
```

Soulution 6:

```
To find out why space in the log cannot be reused, see the log_reuse_wait_desc column in
sys.databases
```

Error 7

"The temp directory in chart handler configuration is not accessible" in Microsoft Chart Control while creating charts.

Solution 7

And what it is when you look in to your web.config you will find this TAG in the appSettings:

<add key="ChartImageHandler" value="Storage=file;Timeout=20;Url=~/tempImages/;"/>

When you drag the Chart control into design mode it adds the tag above to the web.config and you shoud have a folder called TempImages. You get the error when its trying to right or modify that folder and all you need is to give the folder the right access.

On my Vista machine i just gave COMPNAME\Users write and modify access.
But i reckon you mght need to give IIS_WPG(COMP_NAME\IIS_WPG) access too maybe on other OS.

Error 8

If a website is down ater an upgrade, it could be an App_Offline.htm problem

Solution blog website: http://weblogs.asp.net/scottgu/archive/2005/10/06/426755.aspx

Solution 8: Once you remove the app_offline.htm file, the next request into the application will cause ASP.NET to load the application and thewebsite will continue working as expected.

Appendix 44 – Gantt Chart of Project

METHODOLOGY AND TIMELINE

Detailed tasks		*Background tasks*	
1	Literature review	18 - 20	Write-ups and tests
2 – 4	Creating Database structure and checking data	21 - 24	Consultations
quality		25	Milestones
5 -7	Design user interface		
8 – 12	Writing and testing codes for Data Analysis		
13 – 15	Extension of database		
16 – 17	Documentation		

	Months	0-4	5 - 8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40	41-44	45-48
	Tasks												
1	Acquire skills from library & internet												
2	Study architecture of database & webpage												
3	Restructure and cleanup database and website												
4	Update database												
5	Create gene overview page and hyperlinks												
6	Create graphical pages summarizing data												
7	Design webpage for monitoring visitors												
8	Import and process data for association studies												
9	Design webpage to view GWA studies												
10	Import and process data for linkage studies												
11	Design webpage to link groups working on ALS												
12	Combination of Linkage and Association studies												
13	Extend database and webpage to include other species												
14	Write other operational codes												
15	Tidying up database and webpages												
16	ALSOD Documentation												
17	Writing thesis												
18	Testing												
19	Delays												
20	Keep log book												
21	Communicating progress to supervisors												
22	Consultations												
23	Learning bioinformatic tools like R, Plink, Perl etc												
24	Attending meetings, seminars, lectures, conferences												
25	MILESTONES		M1		M2				M3				M4
	Detailed tasks												
	Background tasks												

M1- Completion of database restructure and updates
M2 - Completion of association and linkage studies
M3 - Completion of database extension
M4 - Completion of report writing

Appendix 45 – Glossary

Amino acids	Amino acids are the building blocks used by cells, including motor neurons, to create proteins. Just as the letters of the alphabet can be combined to form an almost endless variety of words, amino acids can be linked together to form a vast variety of proteins.
Atrophy	It is wasting and shrinkage of tissue.
Axon	It is the long, hairlike extension of a nerve cell that carries a message to the next nerve cell.
Biomarker	A biomarker is a substance that can be used to diagnose or monitor disease. Currently, there is no biomarker available to diagnose or monitor ALS. To learn about efforts to identify biomarkers for ALS
Bulbar onset	This refers to the type of ALS where initial symptoms appear in the face and neck, such as difficulty swallowing or forming words.
Cerebellum	It is a large, two-halved structure (hemispheres) located in the lower part of the brain that's responsible for the coordination of movement and balance.
Clinical onset	Clinical onset refers to the time at which signs or symptoms of a disease first appears.
Definite ALS	If loss of upper and lower motor neurons is detected in three or more regions of the body, the disease is diagnosed as definite ALS.
Dementia	The progressive decline in cognitive function due to damage or disease in the brain. Memory, attention, language, and problem solving ability may be impaired. Dementia can also result in confusion. Dementia occurs due to the degeneration of the cortex of

	the brain. Some people with ALS also experience dementia.
Dysarthria	Dysarthria is impaired speech and language due to weakness or stiffness in the muscles used for speaking
Dysphagia	Dysphagia is difficulty in swallowing
Fasciculation	A localized, uncontrolled, uncoordinated involuntary twitching of a single muscle group. Fasciculations could be benign or indicative of disease. Fasciculations are often described as one of the first symptoms of ALS.
Frontotemporal Dementia	Also called frontotemporal lobular degeneration (FTLD), frontotemporal dementia (FTD) is a term that refers to a number of disorders that occur due to shrinkage of the frontal and temporal lobes of the brain. Symptoms include executive dysfunction (difficulties in critical thinking and problem solving), language/speech deficits and behavioral problems. Certain people with ALS also have FTD.
HumVar	HumVar is preferred model for diagnostics of Mendelian diseases which requires distinguishing mutations with drastic effects from all the remaining human variation, including abundant mildly deleterious alleles
HumDiv	HumDiv is preferred model for evaluating rare alleles, dense mapping of regions identified by genome-wide association studies, and analysis of natural selection.
Hyperreflexia	hyperreflexia is excessive response of muscle reflexes when a normal stimulus is applied.
Hyporeflexia	hyporeflexia is weak or absent muscle response when a normal stimulus is applied.

in vivo	In vivo refers to laboratory studies performed within a whole, living organism.
Incidence	The occurrence of new cases of a condition. Incidence is commonly measured in new cases per 1,000 (or 100,000) of population at risk, per year. The incidence of ALS typically varies between 1 and 4 diagnoses per 100,000 of populations per year in the U.S.
Limb Onset	Refers to the type of ALS where initial symptoms appear in the limbs. It is the most common form of ALS, the other type being bulbar-onset.
Lower motor neuron	These are the motor neurons connecting the brainstem and spinal cord to muscle fibers, bringing the nerve impulses from the upper motor neurons out to the muscles. In ALS, the motor neurons degenerate or die, ceasing to send messages to muscles. Unable to function, the muscles gradually weaken resulting in paralysis.
Magnetic Resonance Imaging (MRI)	MRI is a medical imaging technique used to visualize the integrity of specific structures and tissues in the body. It has much greater soft tissue contrast than computed tomography (CT) making it especially useful in neurological, musculoskeletal, cardiovascular, and oncological imaging. Scientists hope to use MRI to diagnose ALS more quickly.
Multiple Sclerosis	Multiple Sclerosis (MS) - a disease of the central nervous system that is unpredictable. MS can be relatively benign, disabling, or devastating, leaving the patient unable to speak, walk, or write.
Neuron	Neuron is a cell specialized to conduct and generate electrical impulses and to carry information from one part of the brain to

	another.
Parkinson's disease (PD)	Parkinson's disease (PD) is a neurodegenerative brain disease that results in shaking (tremors) and muscle stiffness. Other symptoms include trouble walking or maintaining balance.The disease is thought to occur due to loss of dopamine-producing neurons in the brain.
Phenotype	A phenotype refers to everything observable about a living organism. ALS is often a described as a disease with variable phenotypes because the symptoms of the disease can differ between individual patients.
positron emission tomography (PET) scan	A computer-based imaging technique that provides a picture of the brain's activity rather than its structure. The technique detects levels of injected glucose labeled with a radioactive tracer.
primary lateral sclerosis (PLS)	A progressive neurological disease in which the upper motor nerve cells (neurons) deteriorate. If the lower motor neurons are not affected within two years, the disease usually remains a pure upper motor neuron disease. This is the rarest of all forms of ALS.
progressive bulbar palsy (PBP)	A condition that begins with difficulties in speaking, chewing and swallowing due to lower motor nerve cell (neuron) deterioration. This disorder affects about 25 percent of all people with ALS.
progressive muscular atrophy (PMA)	A progressive neurological disease in which the lower motor nerve cells (neurons) deteriorate. If the upper motor neurons are unaffected within two years, the disease usually remains a pure lower motor neuron disease.

Spinal cord	The spinal cord is the part of the central nervous system that extends from the base of the skull through the lower back. It is continuous with the brain stem and encased in a triple sheath of membranes. The spinal cord is typically 15 to 17 inches long and contains 33 vertebrae and 31 pairs of nerves. The spinal cord enables the brain to communicate with the rest of the body.
spinal muscular atrophy (SMA)	A hereditary neurological disease in which only the lower motor nerve cells are affected.
Upper motor neurons	Upper motor neurons (UMNs) originate in the motor cortex of the brain. Upper motor neurons enable certain movements including walking and chewing food.

Appendix 46 - Acronyms

PC - Personal Computer

NIH - National Institutes of Health

CODASYL – Conference on Data System Language

DBMS DataBase Management System

IMS - Information Management System

ORDBMS – Object Relational Database Management System

XML - Extensible Markup Language

API – Application Programming Interfaces

KDD - Knowledge Discovery in Databases

OMIM – Online Mendelian Inheritance in Man

MND – Motor Neuron Disease

ALSA – Amyotrophic Lateral Sclerosis Association

UMN – Upper Motor Neuron

LMN – Lower Motor Neuron

NCBI - National Center for Biotechnology Information

UCSC – University of California Santa Cruz

KEGG – Kyoto Encyclopedia of Genes and Genomes

UNIPROT - Universal Protein

GWAS – Genome Wide Association Study

RSS – Rich Site Summary or Really Simple Syndication

BOAA - β-*N*-Oxalylamino-l-alanine

Appendix 47 - Resources

Scientific websites

1000 Genomes project - www.1000genomes.org/

Exome Variant Server database - http://evs.gs.washington.edu/EVS/

Database of single-nucleotide polymorphisms (dbSNP) - http://www.ncbi.nlm.nih.gov/SNP/

ExonPrimer - http://ihg.gsf.de/ihg/ExonPrimer.html

NCBI37/hg19 assembly - http://genome.ucsc.edu/

NHLBI GO Exome Sequencing Project (NHLBI-ESP) - https://esp.gs.washington.edu/drupal

Online Mendelian Inheritance in Man (OMIM) - http://www.omim.org/

Personal genome databases - http://www.sequenceontology.org/resources/10Gen.html

PLINK algorithm - http://pngu.mgh.harvard.edu/purcell/plink/

PolyPhen-2 - http://genetics.bwh.harvard.edu/pph2/

Refseq - http://www.ncbi.nlm.nih.gov/projects/RefSeq/

UCSC Human Genome Browser - http://genome.ucsc.edu/

NCBI - http://www.ncbi.nlm.nih.gov/gene

Gene cards - http://www.genecards.org/

ExPASy Bioinformatics Resource Portal - http://www.expasy.org/

UniProt - http://www.uniprot.org/

iHop - http://www.ihop-net.org/UniPub/iHOP/

KEGG ALS pathway - http://www.genome.jp/kegg-bin/show_pathway?hsa05014

Genetic Testing Registry - http://www.ncbi.nlm.nih.gov/gtr/

NCBI Gene Review - http://www.ncbi.nlm.nih.gov/books/NBK1450/

Neuromuscular Disease Center - http://neuromuscular.wustl.edu/index.html

Mouse Genome Informatics - http://www.informatics.jax.org/

Gene Ontology - http://www.geneontology.org/

Ensembl - http://www.ensembl.org/index.html

WikiGenes - http://www.wikigenes.org/

ALSmutationDatabase https://reseq.biosciencedbc.jp/resequence/SearchDisease.do?targetId=1

GWAS Phenomap Project – http://www.gwascentral.org/gwasphenomap

ALSGene – http://www.alsgene.org

Bioinformatics tools

GeneMANIA - http://genemania.org/

NetPhos 2.0 Server - http://www.cbs.dtu.dk/services/NetPhos/

Bioinformatics resource portal - http://bioinformaticsweb.net/

Displaying your own Annotations in the Genome Browser - http://genome.cse.ucsc.edu/goldenPath/help/customTrack.html

Human Variation SNP ids - http://www.ncbi.nlm.nih.gov/projects/SNP/tranSNP/tranSNP.cgi

PANTHER – http://www.pantherdb.org/tools/csnpScoreForm.jsp

SIFT - http://sift.jcvi.org/www/SIFT_BLink_submit.html

POLYPHEN - http://genetics.bwh.harvard.edu/pph2/index.shtml

Tutorial websites

Introduction to ASP.Net Tutorial - http://asp.net-tutorials.com/basics/introduction/

Android Development Tutorial - http://www.vogella.com/tutorials/Android/article.html

Learn C++ - http://www.learncpp.com/

Information about genetics - http://www.info.co.uk/genetics?cb=28&cmp=2036

How to track a website - https://developers.google.com/analytics/devguides/collection/gajs/asyncTracking?csw=1

How to make a widget for a website - http://www.ehow.com/how_2059724_make-widget-website.html

Using Microsoft's Chart Controls - http://www.4guysfromrolla.com/articles/111809-1.aspx

The Chi Square Statistics - http://math.hws.edu/javamath/ryan/ChiSquare.html

Using R - http://rdotnet.codeplex.com/

Available R packages - http://cran.r-project.org/

The R Statistical Language and C#.NET - http://www.codeproject.com/Articles/25819/The-R-Statistical-Language-and-C-NET-Foundations

Redirecting to an ASP.NET Mobile Web Page - http://msdn.microsoft.com/en-US/library/fhhycabe(v=vs.80)

Blackberry Developer - https://developer.blackberry.com/

ASP.NET MVC 4 Mobile Features - http://www.asp.net/mvc/tutorials/mvc-4/aspnet-mvc-4-mobile-features

Introduction of Mobile Controls - http://www.codeproject.com/Articles/17842/Introduction-of-Mobile-Controls-Available-in-ASP-N

Mobile Apps & Sites with ASP.NET - http://www.asp.net/mobile

ASP.NET IsMobileDevice and how it works - http://ngeor.net/2011/03/asp-net-ismobiledevice-and-how-it-works/

How to develop Android apps - http://reviews.cnet.co.uk/mobile-apps/how-to-develop-android-apps-50008626/

Expanding/Collapsing GridView Rows - http://www.codeproject.com/Articles/25858/Expanding-Collapsing-GridView-Rows

Server 2005 – Populating the Database - http://www.exforsys.com/tutorials/sql-server-2005/populating-the-sql-server-database.html

ORDER BY statement - http://www.1keydata.com/sql/sqlorderby.html

Ajax Control Toolkit Tutorials - http://www.asp.net/ajaxlibrary/act_tutorials.ashx

Ajax, IE and Firefox! - http://powerdream5.wordpress.com/2007/10/11/ajax-ie-and-firefox/

Digging More intoSQL Server 2000 Using Client-side Javascript - http://www.devarticles.com/c/a/JavaScript/Digging-More-into-SQL-Server-2000-Using-Clientside-JavaScript/

Google Developers - https://developers.google.com/maps/?csw=1

Google Maps via ASP.NET/SQL Server Tutorial - http://forums.asp.net/t/1286639.aspx

How to embed almost anything in your website - http://www.labnol.org/internet/how-to-embed-in-html-webpages/6365/

Google Geomap Visualization API in ASP.Net - http://www.aspsnippets.com/Articles/Google-Geomap-Visualization-API-in-ASP.Net.aspx

Google Maps and ASP.NET - http://be.sys-con.com/node/171162?page=0%2C1

Upgrading Your Google Maps JavascriptApplication to v3 - https://developers.google.com/maps/documentation/javascript/v2/v2tov3?csw=1

Google Maps API Tutorial - http://econym.org.uk/gmap/

Chart Gallery – Google Charts - https://developers.google.com/chart/interactive/docs/gallery?csw=1

SQL Programming - http://publib.boulder.ibm.com/infocenter/iseries/v5r3/index.jsp?topic=%2Fsqlp%2Frbafydelete.htm

Non Scientific websites

Google Earth API - https://developers.google.com/earth/

Java for Window - http://java.com/en/download/win8.jsp

Wikipedia - http://en.wikipedia.org/wiki/Main_Page

WolframAlpha - http://www.wolframalpha.com/

Blogs on Web development

How to create a mobile version of a website http://forums.asp.net/t/1470188.aspx?how+to+create+a+mobile+version+of+a+website+

Fix: Facebook Like Button Not Working - http://www.allwebmaster.com/fix-facebook-like-button-not-working/

Insert Records using TextBox to SQL Database - http://www.bigresource.com/MS_SQL--Insert-Records-Using-from-Text-box-to-SQL-database--P5dXIRD8.html

How to insert data in database using Textboxes - http://www.querycat.com/question/d01e1d84babbb71c2b2385f7a8bde242

Call a perl script using ASP.NET - http://www.asp.net/ajaxlibrary/act_tutorials.ashx

TSQL Text File Output - http://www.dbforums.com/microsoft-sql-server/1607213-tsql-text-file-output.html

How do I send a database query to a text file - http://databases.aspfaq.com/database/how-do-i-send-a-database-query-to-a-text-file.html

Social Media Accounts

Facebook – https://www.facebook.com/pages/ALSoD/307667685943735

Twitter - https://twitter.com/ALSoD_Database